PROCEEDINGS

OF THE SOCIETY OF PHOTO-OPTICAL INSTRUMENTATION ENGINEERS

Volume 228

Active Optical Devices and Applications

William J. Cuneo, Jr.
Editor

April 10-11, 1980
Washington, D.C.

Accession Number: 3868

Publication Date: Apr 11, 1980

Title: Active Optical Devices and Applications, Proceedings of the Society of Photo-Optical Instrumentation Engineers, April 10-11, 1980, Washington, DC

Personal Author: Cuneo, W.J.

Corporate Author Or Publisher: Society of Photo-Optical Instrumentation Engineers, PO Box 10, Belingh Report Number: Volume 228 Report Number Assigned by Contract Monitor: SLL 84-U-98

Comments on Document: Archive, RRI, DEW

Descriptors, Keywords: Active Optical Device Application Large Optics Adaptive Technology Wavefront Sensor Deformable Mirror Performance Coupling Parameter NASA Infrared Submillimeter Astronomy Astrophysics Space Telescope

Pages: 00156

Cataloged Date: Nov 20, 1992

Document Type: HC

Number of Copies In Library: 000001

Record ID: 25159

Source of Document: DEW

Proceedings of the Society of Photo-Optical Instrumentation Engineers

Volume 228

Active Optical Devices and Applications

William J. Cuneo, Jr.
Editor

April 10–11, 1980
Washington, D.C.

Published by
The Society of Photo-Optical Instrumentation Engineers
P.O. Box 10, Bellingham, Washington U.S.A. 98225
206/676-3290

The papers appearing in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and have not been reviewed, refereed, edited, or copyedited. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by the Society of Photo-Optical Instrumentation Engineers.

Please use the following format to cite material from this book:

Author(s), "Title of Paper," *Active optical devices and applications,* Proc. Soc. Photo-Opt. Instr. Eng. 228, page numbers (1980).

Library of Congress Catalog Card No.: 80-50730

ISBN 0-89252-257-7

Printed in the United States of America.

ACTIVE OPTICAL DEVICES AND APPLICATIONS

Volume 228

Contents

Seminar Committee

ACTIVE OPTICAL DEVICES AND APPLICATIONS

Volume 228

Chairman
William J. Cuneo, Jr.
United States Air Force
NASA Headquarters

Co-Chairmen
Ronald F. Prater
United States Air Force
Defense Advanced Research Projects Agency

Thomas G. Pitts
Rome Air Development Center

Charles O. Jones
NASA Marshall Space Flight Center

Chairman Session 1—Technology for Large Adaptive Optics I
Ronald F. Prater
United States Air Force

Chairman Session 2—Technology for Large Adaptive Optics II
Thomas G. Pitts
Rome Air Development Center

Chairman Session 3—Potential NASA Applications of Large Adaptive Optics
Charles O. Jones
NASA Marshall Space Flight Center

ACTIVE OPTICAL DEVICES AND APPLICATIONS

Volume 228

INTRODUCTION

This seminar was organized to bring together the people in DoD and NASA who are developing and applying the technology of active optics for use in space.

In the first two sessions Col. Ronald Prater of the Defense Advanced Research Projects Agency and Mr. Thomas Pitts of the USAF Rome Air Development Center organized a presentation of selected results from the DoD sponsored work on the components for active optics. In the third session Mr. Charles Jones of the NASA Marshall Space Flight Center organized a presentation of the possible NASA applications of large active optics. In the fourth session we discussed in open forum the priorities among the NASA applications and the means for coordination with DoD. The NASA priorities which emerged were (1) an x-ray telescope of 1.2 meter aperture; (2) a 10 to 30 meter diameter telescope for diffraction limited imaging at 30 microns and longer wavelengths; and (3) a 10 to 100 meter diameter thinned aperture optical wavelength telescope that could build two-dimensional resolution by summation of intensity images from a rotating line array of apertures. Treat these priorities as tentative: further deliberation among involved parties is necessary.

This SPIE seminar was especially timely because NASA today has a resurgence of interest in large optics, stimulated by the astronomers in the Office of Space Science. Many years ago, NASA pioneered in active optics by funding at Perkin-Elmer from 1964 to 1970 a three-segment mirror and then a 58 actuator 30 inch shell mirror which achieved a visually spectacular two diffraction ring performance. After 1970 NASA concentrated on getting the space telescope started. Space telescope is now being built and will carry 24 actuators for its primary mirror. Now NASA is looking ahead to the next plateau for astronomical telescopes in space, and this seminar was a step toward that future.

William J. Cuneo, Jr.
United States Air Force
NASA Headquarters

ACTIVE OPTICAL DEVICES AND APPLICATIONS

Volume 228

SESSION 1

TECHNOLOGY FOR LARGE ADAPTIVE OPTICS I

Session Chairman
Ronald F. Prater
United States Air Force
Defense Advanced Research Projects Agency

KEYNOTE ADDRESS

William J. Cuneo, Jr.
United States Air Force, NASA Headquarters

The theme of this Seminar, cooperation between DoD and NASA, reflects a belief that the technology developments of DoD ought to be vigorously made available to NASA whenever possible. NASA seems to have a general image of an agency which conducts spectacular projects in which most people take favorable interest. However, since about 1973 NASA has been moving from the spectacular into a currently more productive direction, the utilization of space. This utilization has many facets. These facets mask the elegance of the potential by breaking it into a detail of apparently lesser projects. But taken as a whole, these projects and their actual or potential results are an important weapon in our nation's strength, to phrase it from a military point of view. The weapons NASA provides are not sharp and well honed like military devices. They reach the spirit. Such weapons are as essential as the sword in the long run. Think about these examples.

NASA science opens the eyes of mankind upon a universe, astounding in its technical and aesthetic majesty. There are few endeavors whose product can touch the spirit of so many people as can the Voyager image of Io floating like a droplet in space across the face of Jupiter.

In a more immediate area, food, the colorful images from Landsat are being successfully interpreted to predict crop production. This use of Landsat imagery has matured to the point where NASA is getting out of the operational business, transitioning activity to NOAA and the Agriculture Department. This ability to foresee food supplies world-wide is an important asset from either a national point of view or the point of view of this planet's people. It is interesting to recall the debate ten years ago on whether Landsat was worth it.

Collection of data related to climate and the environment is another area where there has been impressive progress made through observations with the numerous small and elegant instruments on which many of the attendees at this seminar have worked. A river of data is flowing from these instruments, feeding analysis efforts which are attempting to unravel cause and effect relationships in the climate. This science of climate is too detailed to be as spectacular as a moon landing. But it is likely to be more important. Today some say that we will know in ten years if we are cooking this planet in a carbon dioxide greenhouse. If we find we are, then the data being steadily collected by NASA will become a critical tool in designing a response.

Materials processing is yet another unique area of space utilization. Many multiphase material systems are strongly influenced by the sedimentation or the thermal turbulence effects caused by gravity. Many materials are affected by their terrestrial containers. For these types of systems, the provision of zero-g is likely to be as important a factor as heat is in our conventional material processing experience. We will almost certainly have new products by the year 2000 whose existence depends on substances manufacturable only in space. You will hear few spectacular reports on this. There is a trend to maintain proprietary secrecy as industry teams with NASA to explore this opportunity.

Perhaps the most important potential of NASA is in communications. We have all mostly forgotten how NASA pioneered in the testing of communication satellites in the sixties. Today geosynchronous orbit is like a parking lot with over 60 active civil satellites. You can call anywhere in the nation today for about the same cost because these satellites have removed distance factors from the economics.

NASA is reentering the communication satellite field after a declining period of eight years while the common carriers applied the technology and filled the sky. The goals of the rejuvenated NASA effort are to create technology to open up the use of the 20/30 GHZ portion of the spectrum, to allow more frequency reuse in the crowded lower frequency bands, and to demonstrate limited capacity communication to mobile users. Actually, these goals are mundane because the area is a thicket of regulations, laws, institutions, and least of all, technical problems. But there is a unique light in the air over that thicket—the possibility of powerful communications satellites, short terrestrial tails, and proliferated mediating computers providing highly personalized communications, at affordable cost by definition. Expressed in ultimate terms, communication technology could be advanced by NASA to implement a concept of freedom of access to information and to each other. This is perhaps a contemporary expression of the First Amendment—providing freedom of speech, press, assembly, and petition. As a weapon, such ability to communicate may well be ultimate. With it we can export freedom without a heavy hand. Our present adversaries seem to be able to too soon match our military hardware and perhaps out produce us during peacetime. But they have no effective response to freedom and its fruits in the ongoing battle for men's minds.

Over the past two centuries mankind has enjoyed or endured the machine revolution, the transportation revolution, the energy transmission revolution, and the medical revolution. We are now in the midst of the information revolution and the global destruction revolution. Any military man who has committed to fatalistically push the button yearns, at thoughtful times, for an option besides terror and its risks which can still preserve our freedom. In the information revolution may be the seeds of an option. Communications between people, individuals, might lead ultimately to a sufficient empathy between groups so that the fireballs of thermonuclear war don't erupt to settle differences.

How does NASA fit into such a concept? The successful economics of such communication capability almost surely require, among other hardware, powerful communication satellites to move high amounts of information with a minimal terrestrial tail. NASA seems to be the only organization where such possibilities can flourish. Industry cannot take the economic risks. DoD necessarily has to build sharp weapons, not civil communications. By leading the development of such a communication future NASA could perhaps most completely fulfill its legal charge from the Congress to develop space for the benefit of all mankind.

Taken as a whole, NASA is a rich field of opportunities which deserve vigorous pursuit within NASA and support by other agencies, notably DoD. I hope the frequency of joint seminars such as this one continues to increase.

Wavefront sensors and deformable mirrors for visible wavelengths

Noah Bareket
Lockheed Missiles and Space Company, Palo Alto, California
Lawrence T. James
U.S. Air Force, Space Division, Los Angeles, California

Abstract

A review of the present technology status of wavefront sensors and deformable mirrors is presented. As components in adaptive optics systems these elements have been developed for specific requirements. Future applications will require advanced capabilities for visible wavelength operation. Some forecasts of these requirements are provided.

Introduction

Imaging and focusing through the atmosphere have been primary applications for adaptive optics concepts. Many concepts and analyses[1] have been reported in the literature but fewer experimental efforts have been completed. This paper is a brief review of two important components of adaptive optics concepts. The development of wavefront sensors and deformable mirrors has provided hardware prototyping and concept verification. However very few actual systems with operational and field test capability, have been attempted. This lack of system operation inhibits engineering verification and performance measurements. The authors' past involvements in development and analysis, and participation in recent technology reviews have provided the ground work for a review of the two components and an initial consideration for visible wavelength applications. Many concepts and hardware tests of wavefront sensors and deformable mirrors have been for infrared applications. Primarily though breadboard and initial experiments have used visible light for early tests. For future system application with visible light, scale up and upgrades only, would be required. A review of system concepts and components may help future efforts in system design and integration of components.

Wavefront Sensors

The function of the wavefront sensors reviewed here is to provide a real time phase map of an aberrated wavefront. The phase information is processed and is the input to a deformable mirror for correction of aberrations. This is only one of the ways adaptive optical systems can operate[1,2].

The functions of wavefront sensors have been developed for many years for optical testing purposes and it is mainly the temporal requirements which set the two applications apart. Conventional interferometry, which relies on recording interferograms on a film or on a videotape, has been converted into real-time interferometry by heterodyne techniques. It is appropriate, therefore, to review here the principles of heterodyne interferometry. Interferometric sensor concepts are described next, and the application of Hartmann test methods to wavefront sensors is reviewed.

Heterodyne Interferometry

In heterodyne interferometry [3] a frequency shift or phase modulation is introduced

between the two interfering fields. In conventional interferometry, the irradiance at a point $\bar{r}$ can be described as

$$I(\bar{r}) = I_1(\bar{r}) + I_2(\bar{r}) + 2\sqrt{I_1(\bar{r})\, I_2(\bar{r})}\ \gamma \cos\left[\phi(\bar{r})\right]$$

where $I_1 + I_2$ is the irradiance with no interference, γ is the degree of coherence and the optical phase difference $\phi(\bar{r})$ is given by

$$\phi(\bar{r}) = \frac{2\pi}{\lambda} \mathrm{OPD} \qquad (\mathrm{OPD} - \text{optical path difference}).$$

If the optical phase of one of the interfering fields is modulated by (t), the irradiance Z(r) takes the form

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\ \gamma \cos\left[\phi_0(t) + \phi(\bar{r})\right]$$

A linear phase shift $\omega_0 t$ is obtained by introducing a frequency shift between the interfering fields, and the irradiance $I(\bar{r})$ is oscillating at the frequency with a phase $\phi(\bar{r})$. The wavefront information is now encoded into an easily detected signal. The advantages of heterodyne inteferometry are manyfold. Realtime measurements are possible. Today's electronics make it possible to determine the phase of an oscillating signal to high accuracy. The phase measurement is independent of the irradiance distribution across the pupil and the sensitivity of this technique provides high S N R even with low fringe visibility.

Several techniques are available to detect the phase. The zero crossing technique is a common method. The sinusoidal signal is fed into a comparator which changes polarity when the input signal goes through zero. The resulting binary signal starts a counter which counts pulses of a high frequency clock. The reference signal pulse stops the counter and the number of counts is proportional to the phase difference between the signals. Another method is to sample the signal several time during each period and calculate the phase from the sampled data. This method of "sample and convert" can vary in the number of samples taken per period and the algorithm used to recover the phase. Phase sensitive filters or other types of signal processors can be used to enhance the S N R.

The phase modulation can be introduced in various way, as described here.

Moving Mirror. When light bounces off a moving mirror the frequency changes -the well know Doppler effect. The frequency shift is equal to $2v/\lambda$ where v is the mirror velocity component along the direction of incidence (or reflection).

Moving Grating. When a diffraction grating is moved at a velocity v perpendicular to a beam of light the diffracted light is frequency shifted by $2n\pi\, v/d$, where d is the grating constant and n is the order of diffraction.[4]

Acousto-Optic Cell. The acousto-optic interaction is analogous to the moving grating interaction, where $v/d = f$, the frequency of the acoustic wave in the cell.

Electro-Optic Cell. An electro-optic cell changes birefringent characteristics as a function of an applied electricfield. As a result, the phase of linearly polarized light passing through the cell can be made to be directly proportional to the electric field applied.

Rotating Wave-Plate. When circular polarized light passes through a rotating half-wave plate, the frequency of the light is shifted by twice the angular frequency of rotation of the wave plate.[3]

Interferometric Sensors

Self-Referencing Interferometer. This interferometer is a modification

of the point-diffraction interferometer, first described by Smartt and Strong[5]. As depicted in Fig. 1 the wavefront is split into two arms. One beam is focused on a pinhole which is smaller than the Airy disk diameter, a familiar spatial filter. The emerging wavefront is free of aberrations and serves as the reference wave for interference. The phase modulation or frequency shift which is needed for the heterodyne opeation is introduced in the other arm, and in principle any of the methods described above could be used. In another configuration the pinhole is replaced by a reflecting dot on a transparent optical surface, achieving more compact and stable construction.

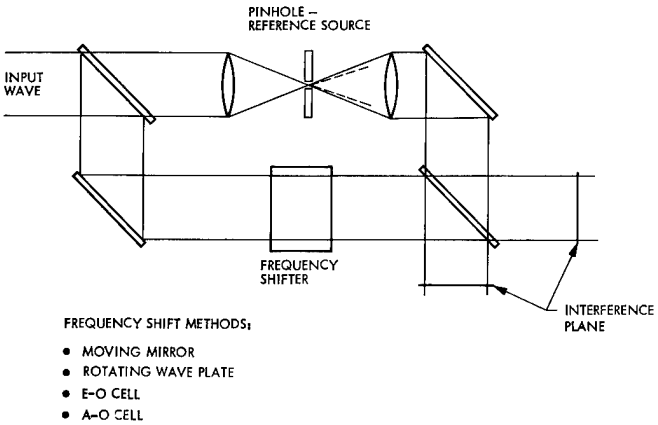


Figure 1. Self-referencing interfermeter

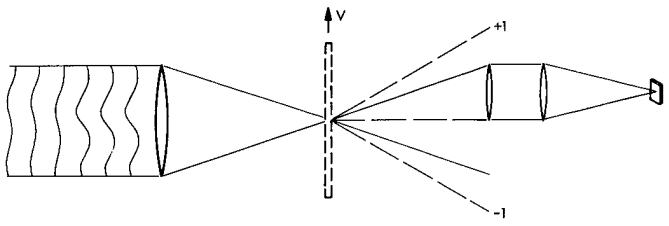


Figure 2. Moving grating lateral shear interferometer

Lateral Shear Inteferometer. In this type of interferometer the wavefront is displaced laterally and interferred with itself, producing an interference pattern which carries informa-tion about the wavefront slope.[6] There are many configurations of interferometers which use this principle, but the one most commonly applied to wavefront sensing is the moving grating interferometer. The principle of a grating shearing interferometer was first described by Ronchi[7] and modified by Wyant[8]. The wavefront is focused on a grating. The light is diffracted at angles which are proportional to the grating frequency and to wavelength. In the most simple configuration the first orders of diffraction are separated and interfere with the zero order only. The shear distance in this case equals half the pupil diameter (Fig. 2). This large amount of shear limits the sensitivity of the interferometer to low spatial frequency wavefront errors. To produce small shear without overlap of the diffraction orders, a double frequency grating is used[8]. To achieve heterodyne operation the grating is moved laterally. Hardy et al described a moving grating shearing interferometer which uses a radial grating, deposited on a rotating disc[9]. This interferometer has several unique features. One optical element, the grating, is used to produced both the wavefront shear and the frequency shift. The amount of shear is proportional to the distance from the center of the radial grating, allowing a convenient way to vary the sensitivity and dynamic range of the interferometer. Finally, overlap of the various diffraction orders is allowed, since the heterodyne frequency can be filtered electronically to allow detection of the zero and first orders interference only.

Hartmann Sensors

The Hartmann test, a classical method of testing large optical elements, can be readily adapted to be used for wavefront sensing. In the classical configuration of the tests, a mask is used to divide the wavefront into subapertures and a lens focuses the sampled wavefront onto a detector plane (Fig. 3a). As in geometrical ray tracing, the position of the light spots in the image plane is related to the slope of the wavefront at the subapertures. The classical Hartmann sampling technique is not efficient. To increase the light efficiency and to avoid overlaping of the sample spots at the focal plane other sampling methods are used. Sampling can be done with a lens array as in Fig. 3b or by an array of wedges, or tilted refractive plates, as in Fig. 3c. A sampling method which is suitable for sampling the output beam of large telescopes is shown in Fig. 3d. The light is sampled by low efficiency gratings which are deposited on the primary mirror. The gratings, which can be generated by holographic techniques, can be made as Fresnel zone gratings and focus the sampled light at a convenient sensor position.

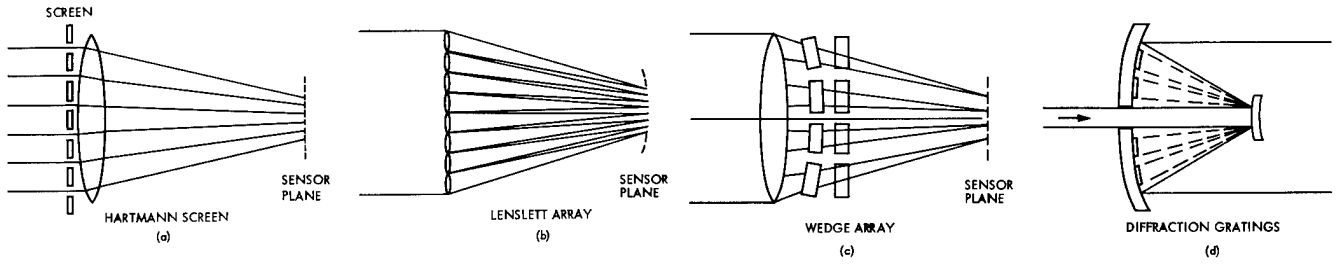


Figure 3. Hartmann wavefront sensors

The relation between the position of the light spots at the focal plane and the wavefront depends on the definition of "spot position". A common definition is the centroid, or center of gravity of the irradiance distribution. Others are the median of the distribution, the peak, or some weighed average of the irradiance. The centroid of the spot is related to the wavefront through the moment theorem, which relates the moments of a function to the derivatives of the fourier transform of that function. Assuming a wavefront distribution $a(x,y) \exp\frac{2\pi j}{\lambda} w(x,y)$, the centroid $(\bar{x},\bar{y})$ of a Hartmann spot is given by

$$\bar{x} = \frac{\iint |a(x,y)|^2 w'(x,y)\,dxdy}{\iint |a(x,y)|^2\,dxdy}, \quad \text{and likewise for } \bar{y}.$$

The integrals are taken over the subaperture area. This particular definition of the spot position gives a measure of the average wavefront slope, weighed by the subaperture irradiance distribution. It is worth noting that estimating the spot position by the centroid algorithm is optimal by the maximum likelihood criteria, when a white Gaussian noise process is assumed.[10]

A Hartmann based wavefront sensor has to track the position of the many sample spots in real time. The various methods can be divided into two categories - image conjugate detection and pupil conjugate detection. The first category includes sensing schemes in which detectors are put in the focal plane or in a relayed image of the focal plane, and the irradiance distribution at this plane is directly analysed. The detection scheme can use, for example, quadrant detectors, large area CCD arrays or scanned detector configurations. The second category performs some kind of optical filtering operation at the focal plane and detects the signals at a pupil conjugate plane. Shearing interferometers, for example belong to this category. Pupil conjugate sensors allow overlap of the Hartmann spots at the focal plane which is sometimes advantageous. They also have better rejection of stray light since each subaperture is imaged onto its corresponding detector.

Image Conjugate Sensors. One of the simplest and most commonly used schemes of position sensing is the quadrant detector. It is usually used as a null detector, sensing the median of the irradiance distribution. When applied to Hartmann spot position sensing it suffers serious drawbacks: limited dynamic range, nonlinearity and poor noise rejection are some of the problems.

The problems associated with quadrant detectors can be overcome by employing large 2D detector arrays to track the spot position. Such an array typically covers more than a single Hartmann spot, and depending on the number of spots and the required spatial and temporal resolution, several arrays can be operated in parallel. Unlike the quadrant detector, where the detector size has to be larger than the spot size, the array detector size is smaller than the spot size.

Once the irradiance distribution is sampled, various algorithms and noise reducing schemes can be implemented. The signal processing problems associated with this type of a Hartmann sensor are identical to the problems associated with imaging trackers, and extensive theoretical and experimental studies have been conducted in this area[10]. Among the algorithms tested are centroid algorithms and edge and corner tracking algorithms.

Threshold gating and correlation techniques have been used to increase the S N R. In this type of sensors there is a tradeoff between the spatial resolution and the frame rate. As the number of spots or the number of pixels per spot increase, the processing time increases. Processing time can be shortened by employing random access arrays, for example charge injection devices. Another alternative is the use of hybrid analog-digital algorithms. The centroid algorithm, for example, can be performed by analog line integration and digital multiplication.

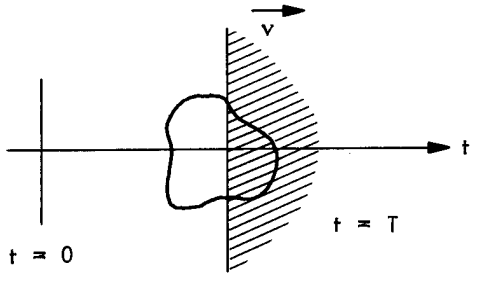
A different approach to position sensing is the use of scanned detectors, rather than staring arrays. Various concepts differ in the detector arrangement, e.g., linear vs. quadrant, and the scanning method. A common requirement for these schemes is the presence of reference beams. The alignment tolerances and stability of scanned systems are less than the requirements of wavefront sensors. By using reference beams which share the same optical path as the measured Hartmann sample beams the mechanical errors are calibrated out. The signal processing can be purely digital, as with a staring array, analog or hybrid.

Pupil Conjugate Sensor. This type of sensor performs an optical filtering operation at the focal plane and detects the irradiance at a plane conjugate to the pupil. Unlike the image conjugate sensors, overlap of the sample spots are allowed at the focal plane, and this reduces the field of view requirements for the sensor. These type of sensors also have better stray light rejection characteristics. The most advanced wavefront sensor concept in this category is the tilt sensitive shearing interferometer. This is a special case of the lateral shear interferometer described before: the amount of shear is adjusted to be half the subaperture size. There is only one resolution point per sampled beam and the detected wavefront slope is thus averaged over the Hartmann subaperture.

Another scheme for a pupil conjugate sensor uses a knife edge to scan the focal plane while the transmitted irradiance is collected on a single detector (Fig 4). It is shown [1] that the centroid of the scanned spot as a function of the scanning velocity v, the scan time T and the detector output V is given by:

$$\bar{x} = vT - \frac{v}{V(T)} \int_0^T V(t)dt.$$

This sensor type suffers from the difficulty of producing the required constant speed scan. When rotational chopping is used, a second order correction is needed, as the scan velocity varies along the radius.



DETECTOR OUTPUT $V(t) \propto \int_0^{vt} dx \int_{-\infty}^{\infty} E(x,y)\, dy$

Figure 4. Centroid detection by image chopping.

Applications in infrared make requirements unique. It is noteworthy that most demonstrations, experiments and diagnostics are conducted with visible light as we mention in the case of wavefront sensors. Many issues and problems have been predicted or discovered during the past years of development. A summary of most of these is provided here.

The mirror concept is an initial consideration. The ultimate use and set of requirements for operation must be as well defined as possible. The characteristics of mirrors as discussed here should be helpful in making design or system integration decisions. The basic mirror design must be compatible with the optical system for which the use is intended, and the available power system.

A control system must be capable of operating the deformation of the mirror. Ease of manufacture and maintenance has been a problem with some designs. Time consuming fabrication processes and risky assembly procedures should be avoided. Future mirror construction programs should consider the tradeoffs of performance for simplicity.

The specification items for mirrors are the same for deformable mirrors with all the additional items associated with active control. The first consideration is most probably size. For small sizes up to say ten centimeters mirror surface deformation requirements may be satisfied by the monolithic block design[18]. Indeed the spatial resolution may require the large number of actuation points. Large sizes up to a few meters may best be constructed using faceplate and actuators. Larger sizes would have to be of segmented design because of near term manufacturing limitations[19]. Along with size, weight may be a factor to consider if a mobile system or a space based system is being designed. Weight impacts materials choices and must be considered along with thermal expansion, material stiffness and fabrication issues. Next the surface quality of the mirror is important. Typical preferred values of RMS surface deviation from desired profile are one twentieth or less of the operating wavelength or wavelength range. Capabilities are improving though and one may expect better quality in the future. Special problems have been encountered in grinding polishing and coating deformable mirror surfaces. Some efforts to use low, zero, or negative lap pressures during grinding and polishing have been successful. Mirror assembly before or after polishing is a design consideration. Assembly can destroy surface quality in case of inadequate manufacturing tolerances. If the resurfacing or repolishing of a mirror requires disassembly this process can be timely and costly. Surface cleaning and maintenance has also been neglected during some design programs. As in all mirror fabrications good diagnostics and evaluations are required.

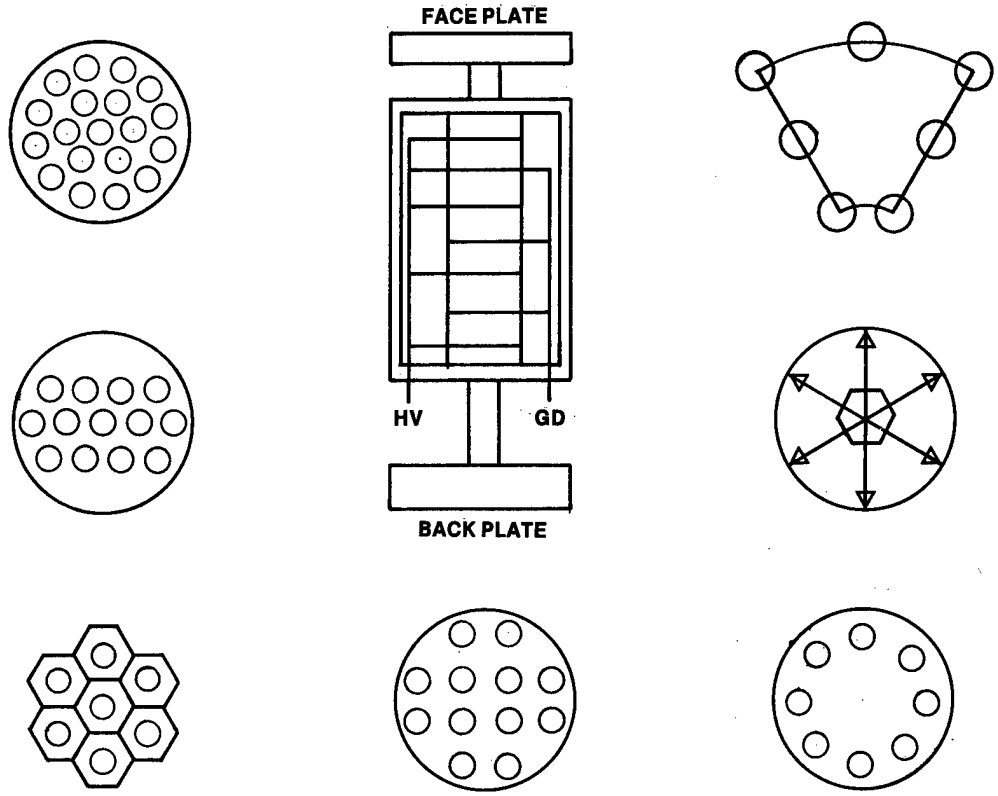


Figure 5. Actuator design and arrays

System Implementation

So far, only a few systems have been reported to demonstrate the speed, resolution and accuracy required of wavefront sensors (as defined here).

A real-time interferometer which is adequate for adaptive optics purposes has been built and tested by ITEK[12]. It is a lateral shear interferometer which uses a radial rotating grating to provide variable shear. Two CCD detector arrays are used to provide x and y axis information. The detector elements integrate the oscillating signal and are read out four or eight times during each period. Simple algorithms are used to calculate the phase of the signal from this data, with a spatial resolution of 24 x 32 pixels. The frame rate is 355 Hz or 710 Hz, depending on the signal sampling rate. Under laboratory conditions the sensor has demonstrated measurement repeatability of 0.004 waves rms. An earlier sensor of the same type has been incorporated into ITEK's Real Time Atmospheric Compensation system which demonstrated closed loop operation.[9]

A wavefront analyzer has been built by Perkin-Elmer[13] to analyze CO_2 laser beams. It was not designed for adaptive optics purposes, and is not light efficient. However, it's spatial resolution (32 x 32) and frame rate (100 Hz) qualify it as a wavefront sensor. The mode of operation is equivalent to a lateral shear interferometer. The wavefront is scanned by a pair of small apertures and, as with a shearing interferometer, the sample points on the wavefront interfere. The interference is being produced in series, hence the low efficiency, and tilt component is superimposed on the interference pattern due to the sampling method.

A geometric type wavefront sensor has been developed by Adaptive Optics Associates.[14] It is of the scanned array type where the Hartmann spots are circularly scanned over corresponding quadrant detectors. A perfect wavefront is injected into the optical train to provide alignment reference. The reference beam is amplitude modulated and the detectors' output is filtered to separate the reference and wavefront information. This system processes the spot position information in parallel and has demonstrated the ability to detect weak signals (S N R = 1) at a measurement rate of 100 Hz and accuracy of 1/15 wave rms.

Deformable Mirror Concepts

The most widely used concept for deformable mirrrors is that of the thin faceplate with attached actuators.[15] Glass or metal faceplates typically of tens of mils thickness are fabricated and an array or pattern of actuators is attached by bonding or brazing. Actuators are of various types discussed below and are generally piston action. The actuator motion or force will cause surface distortions and can permanently deform the faceplates. Backplates are required for actuator mounting. Actuator array patterns (See Figure 5) were designed to provide close positioning of actuator points and matching of surface figures with deformations. The square and hexagonal patterns satisfied close packing for present actuator sizes and designs. The radial patterns were developed to apply the figures derived from the terms of the Zernike polynomial expansion which approximate the Seidel aberrations. The overall mechanical configuration of the faceplate backplate and the actuators impacts the bandwidth of a closed loop active control system. The lowest natural vibration frequency of the configuration will specify the operating limits of automatic control. The faceplate vibration modes such as "drum head" or interactuator areas are generally lowest frequency. Poor actuator designs can have low resonances as well. Some of these problems are not associated with the "monolithic" design[16]. A faceplace is attached to a large (monolithic) block of piezoelectric material. Electrodes are imbedded in the material. A voltage applied to the electrodes produces an electric field to the grounded surface, typically a metal coating on the back of the block but possibly to the metal mirror surface. The piezoelectric effect distorts the block and deforms the mirror surface. The electrode approach can provide closer packing of control points. Segmented mirrors can be simply arrays of these continuous faceplate designs in case of large diameter or can be small segments close packed to minimize diffraction. Individual segments are actuated. Hexagonal areas are favored but round mirrors have been used in the lab to demonstrate feasibility by this author (LJ).

Other concept for deformable mirrors have been concieved and in some cases initial experiments have been conducted. Electrostatic actuation of membranes or thin films has been investigated. Hydraulic pressure actuation of membranes was considered. Advanced concepts have included aerowindows and plasma volumes which have varible and controllable index of refraction.

Design Issues

The designs of deformable mirrors have been in many cases wavelength dependent.

Actuator concepts are a significant part of deformable mirror choices. The considerations include response, extension, and design. Actuator concepts will require continued development. The monolithic block design does not require individual actuator design but has special problems. Imbedding electrodes in the piezoelectic block leaves little ease in replacement or repair of malfunctions. Experience has shown that quality in assembly has precluded any failures.

Influence profiles between electrodes must be characterized. The bonding between actuator block and faceplate is still in development stages. Design trades must consider frequency response, voltage, extension and figure quality as interacting quanities. Typically higher frequency operating materials give shorter extension regardless of design. Individual actuator designs include stacks of piezoelectric materials, electromagnetic devices and hydraulic pushers. Piezoelectric materials were linearly stacked so that alternating electrodes could be attached. Lower voltage drops across multiple discs provided more extension (Fig 5). Actuators have been designed with two axis tilt and piston all in one housing for small elements of a segmented mirror.

Piezoelectric materials have demonstrated characteristics which are not desired and must be compensated. A trade-off has been extension for hysteresis effects. In open loop control registration is lost for higher hysteresis effect materials. Lower hysteresis materials show shorter extension. Not a lot of materials effort has been invested to solve these problems. Electromagnetic actuators have been used for tilt - only mirrors. Hydraulic actuators pose a new set of problems such as fluid handling and acoustic vibration. Although designs exist only prototypes have been demonstrated. The signal feed thru mechanism for the hydraulic actuator is special design problem however all of actuators need consideration in this area. All of the actuator designs are difficult to manufacture, assemble, and repair.

The electronics systems used with deformable mirrors include actuator drive as well as control systems. The control functions must be directed by the system to drive the mirror profile to a certain configuration. Open loop control schemes have not been as preferable as closed loop. The design considerations are frequency response and integration with hardware. Hopefully the electronics will not be the overall limitation to system bandwidth.

The mechanical resonances must be the limitations and electronically driven functions will fall within these limits. Bandwidths in actual systems have been lower than those given below for various concepts. (Table II)

Mirror Performance

A number of deformable mirrors which have been developed for specific systems have not been well characterized in the literature. Some performance numbers are listed in table II. These values are typical for the various concepts and implementations. Three types of designs are reported. The response frequencies are measured response below the major natural resonances of the mirror structure and assembly. Typical closed loop active control has not approached these frequencies. The extension values have been closer to requirements. Control systems to drive these deformable mirrors have not been implemented to full capability. (Table II)

Future Requirements

Past efforts have uncovered many problems and provided some solutions. A review of these efforts allows some speculation for future configurations. Imaging and focusing systems operating with visible wavelength will require more actuators and better surface quality. The monolithic design appears to be promising for this application up to nominal diameters.[18,20] For larger mirror edge actuation and segments seem to be good choices.[19]

Table II

DEFORMABLE MIRROR PERFORMANCE

CONCEPT	SIZE	THROW	RESPONSE	ACTUATORS
METAL FACEPLATE PZT ARRAY	20 CM	$\pm .4\,\mu$ TO $\pm 35\,\mu$	3.5-30 KHZ 200 HZ (CLOSED LOOP)	50-70 SQ OR RADIAL ARRAY
METAL FACEPLATE PZT EDGE (REF 17)	10 CM	$\pm 9\,\mu$	200 HZ	6
MONOLITHIC ELECTRODES	23 CM	$\pm 9\,\mu$	2 KHZ	37 RADIAL ARRAY

Conclusion

How does the technology reviewed here compare with the requirements of adaptive optical systems operating at the visible part of the spectrum? For improving the resolution of astronomical imagery, the spatial resolution, frame rate and accuracy of existing components allow the construction of functioning systems. For other applications, however, the state-of-the-art is far from satisfactory. Consider the potential application of using visible wavelength lasers for global communication systems. Assume that it is necessary to transmit useful laser energy to a geosynchronous satellite, a distance of 4×10^7 meters away. A first order approximation of the number of correcting elements required for wavefront control is $(D/r_o)^2$ where D is the telescope diameter and r_o is the atmospheric coherence diameter. The telescope is assumed to be 4-5 meters in diameter, and r_o is estimate to range between 7 cm and 10 cm. Thus, up to 5000 correction channels might be needed for the active element and a comparable spatial resolution is required for the wavefront sensor. The temporal requirements are not trivial. A turbulence bandwidth of 50-150 Hz requires a control bandwidth on the order of 1 KHz and a wavefront measurement rate which is several times faster. While there are no fundamental difficulties in upgrading the existing wavefront sensing technology to match these requirements, it will be difficult to achieve the high actuator density with today's deformable mirror technology. One solution may be the actuation of the largest mirror in the telescope - the primary mirror where the density will be minimized. Another solution may do away wth deformable mirrors and wavefront sensors altogether by using nonlinear phase conjugation methods.[21]

References

1. J. W. Hardy, "Active Optics: A New Technology for the Control of Light," Proc. IEEE Vol 66, page 651 (1978)

2. J. E. Pearson, R. H. Freeman and H. C. Reynolds, Jr., "Adaptive Optical Techniques for Wave-Front Correction" in Applied Optics and Optical Engineering, Vol. 7, R. R. Shannon and J. C. Wyant, Eds. (Academic Press, New York, 1979).

3. J. H. Bruning "Fringe Scanning Interferometers," in Optical Shop Testing, D. Malacara, Ed., (Wiley, New York, 1978). Includes more references on heterodyne interferometry.

4. R. H. Stevenson, "Optical Frequency Shifting by Means of a Rotating Diffraction Grating", Appl. Opt. Vol 9, page 649 (1970),

5. R. N. Smartt and J. Strong, "Point Diffraction Interferometer," J. Opt. Soc. Am. Vol 62, page 737 (1978)

6. M.V.R.K. Murty, "Lateral Shearing Interferometers", Optical Shop Testing, D. Malacara, Ed. (Wiley, New York, 1978).

7. V. Ronchi, "Forty Years of History of a Grating Interferometer," Appl. Opt. Vol 3, page 93 (1964).

8. J. C. Wyant, "Double Frequency Grating Lateral Shear Interferometer," Appl. Opt. Vol 12, page 2057 (1973).

9. J. W. Hardy, J. E. Lefebore and C. L. Koliopoulos, "Real Time Atmospheric Compensation," J. Opt. Soc. Am., Vol 67, page 360, (1977)

10. G. Gerson and A. Rue, "Tracking Systems", The Infrared Handbook, W. L. Wolfe and G. J. Zissis, Eds., (Office of Naval Research, Washington, D.C., 1979)

11. N. Bareket and W. L. Wolfe, "Image Chopping Techniques for Fast Measurement of Irradiance Distribution Parameter," Appl. Opt. Vol 18, page 389 (1979).

12. K. D. Stumpf, "Real-Time Interferometer," Opt. Eng. Vol 6, page 648 (1979)

13. J. Van Workum, J. A. Plascyk and M. L. Skolnik, "Laser Wavefront Analyzer for Diagnosing High-Energy Lasers," Adaptive Optical Components, Proc. SPIE, Vol 141, page 58 (1978).

14. L. E. Schmutz, J. K. Bowker, J. Feinleib and S. Tubbs, "Integrated Imaging Irradiance (I) Sensor: A New Method of Real-Time Wavefront Measuration," Adaptive Optical Components II, Proc. SPIE, Vol. 179, page 76 (1979)

15. J. E. Harvey and G.M. Callahan," Wavefront Error Compensation Capabilities of Multi-Actuator Deformable Mirrors," Adaptive Optical Components Proceedings of SPIE, Vol. 141, p 50 (1978) Washington D.C.

16. J. Feinleib, S. G. Lipson, and P. F. Cone "Monolithic Piezoelectric Mirror for Wavefront Correction", Applied Physics Letters Vol 25, page 311 (1 Sept 74)

17. Proceedings of the Society of Photo-optical Instrumentation Engineers, Vol 179, Adaptive Optical Components II Technical Symposium East April 1979 Washington D.C.

18. Aldrich, R. E., J. H. Everson and N. P. Albertinetti "High Bandwidth, Fine Resolution Deformable Mirror Design," AFWL-TR-79-5, March 1980 (distribution unlimited) FTR Contract F29601-78-C-0060

19. D. M. Aspinwall and T. J. Karr "Improved Figure Control with Edge Applications of Forces and Moments", Active Optical Devices and Applications, proceedings of the SPIE, Vol 228 Washington D.C. (1980)

20. J. H. Everson, M. Cone and J. Kenemuth "Device Parameters and Optical Peformance of a Stacked Actuator Deformable Mirror" Active Optical Devices and Applications, proceedings of the SPIE Vol 228 Washington D.C. (1980)

21. V. Wang, "Non-Linear Optical Phase Conjugation for Laser Systems" Optical Eng. (USA) Vol 1 No. 3, p. 267 (May - June 1978)

Experimental performance of the I^3 wavefront sensor for closed-loop adaptive optics

L. E. Schmutz, J. K. Bowker, J. Feinleib, S. N. Landon, S. J. Tubbs
Adaptive Optics Associates, Inc.
12 Prentiss Street, Cambridge, Massachusetts 02140

Abstract

Analytical and experimentally determined performance characteristics of the Integrated Imaging Irradiance (I^3) Sensor are presented with emphasis on features relevant to closed-loop operation. Hardware components which were used to develop data on ultimate S/N performance of the sensor are described. Experiments performed on the I^3 Sensor breadboard under computer control in which optical tilt variations in an input wavefront, induced either by an electronically controlled jitter mirror or by artificially generated atmospheric turbulence, are compensated using a Type 1 digital servo, are detailed. Other sensor features reviewed include operation in the presence of partial detector failure, and with either negative or positive contrast sources.

Introduction

High speed optical feedback systems have established their importance in a wide variety of applications which range from photo-mask alignment systems, auto-focussing cameras and interferometric control of machine tools, to laser amplitude, frequency and phase front control, tracking systems, and, of course, real-time image compensation. The successful use of the optical feedback concept in so many problem areas indicates the maturity of the technology. Adaptive optical devices are coming into wider use as system components as their capabilities and limitations become well understood.

The I^3 Sensor[1,2,3] is a fast dynamic optical tilt sensor capable of operating with extremely low light levels. Originally developed as a many-subaperture dynamic wavefront analyzer used in phase conjugate adaptive image compensation, it can also be used in single channel form as an optimal tracker. Extensive characterization of the I^3 Sensor has been carried out not only analytically and by numerical simulation but through considerable experimental investigation as well. In this way, features specifically relevant to use of the I^3 Sensor as a feedback system component have been determined, and these features will be emphasized below.

Sensor operation

The structure and operation of the I^3 Sensor has been presented in detail elsewhere[2,3,4]. It is derived from the Hartmann test, which determines the mean tilt of a source wavefront by focussing the wavefront to a diffraction spot and measuring the centroid of the spot position relative to some reference point. In the I^3 Sensor, that reference point is the focal spot of a locally generated plane wave which has been added to the source wavefront. Determining the difference between the two spot positions therefore measures the difference between the source and reference wavefront tilts.

In the I^3 Sensor, spot position detection is done with quadrant detectors. A conically scanned nutation mirror causes the source and reference spots to follow circular paths on the quadrant detector so that each quadrant output rises or falls as a spot enters or leaves it. In this way, the input optical signal is efficiently placed on an AC carrier determined by the nutation rate. Taken together, the four quadrant detector waveforms unambiguously encode any focal spot position and hence any wavefront tilt within the field of view of the instrument.

This basic structure implies an inherently stable measuring system since it is differential, combined with an efficient optical modulation technique. More careful study has revealed a number of interesting sensor properties which can be practically exploited.

Graceful degradation

The data processing algorithms which are of use in extracting spot position information from the detector waveforms share a number of useful features. For example, although four detectors are required for best S/N performance, correct tilt sensing is still obtained if a detector fails. Figure 1 shows the effective coordinate system distortion in a source-spot only measurement that results from a quadrant four (lower right) detector failure.

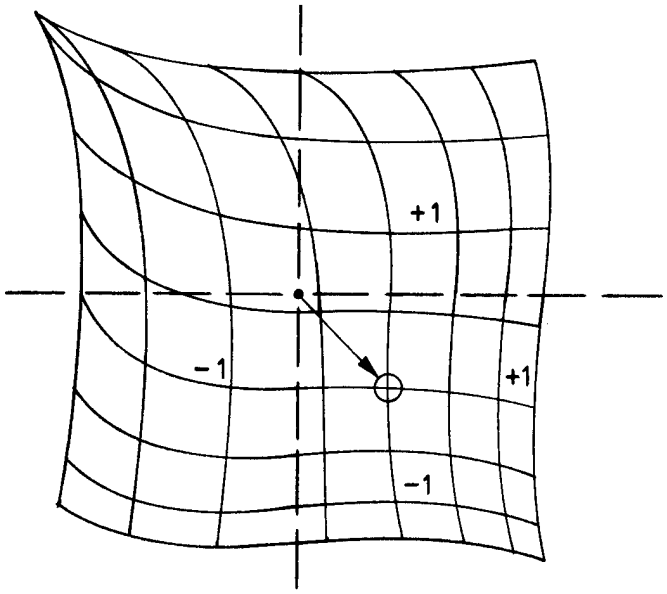


Figure 1. Apparent source coordinate system resulting from a failure of the quadrant 4 detector. The coordinate system is distorted, with the apparent origin shifted towards the fourth quadrant. This simulation was prepared for spot radius of unit size, a nutation radius of unity, and a quad detector extent of 4.3 units.

This mapping was obtained using a numerical simulation of the detection process. The major consequence of the detector loss is an apparent shift in the location of the origin towards the dead quadrant, and a warping of the cartesian geometry which becomes pronounced near the quad cell edges. However, the mapping at no point becomes double-valued. This means that when the equivalent position map of the reference spot is subtracted from this one (which is the final step in the I^3 process) the true origin is restored, since an identical distortion will occur for both source and reference spots. This implies that in a closed-loop system, an accurate null will be retained in the presence of a detector failure with a reduction in S/N ratio of $\sqrt{3/4}$. In fact, the I^3 Sensor will continue to operate satisfactorily if two detectors fail so long as these are located in diagonally opposite quadrants. More generally, this implies that use of a reference relaxes the necessity of good matching between the detectivities of the quadrant detectors for accurate performance.

This result is predicted not only by simulation but verified experimentally using a laboratory version of the I^3 sensor.

Hardware systems

Two important experimental systems have been developed for evaluating I^3 Sensor performance: 1) an electronic emulator has been built which generates the waveforms expected from the I^3 quadrant detector under a large variety of possible circumstances, but with very precise control. It implements a model of the I^3 Sensing process based on a square-shaped, uniformly illuminated focal spot. Nutation radii of from 0 to $\sqrt{2}$ W, where W is the spot width, can be emulated at nutation rates of 20 Hz - 20 kHz. Displacements, both static and dynamic, of source and reference spots from the quad cell center may be generated, and the resulting waveforms for each quadrant combined with a controlled amoung of uncorrelated noise. This system has been highly useful in debugging the I^3 hardware and software processors, as well as in exploring their performance limits. 2) A hardware breadboard has been constructed on which many kinds of wavefront sensing experiments may be performed. Described in detail in Figure 2, the breadboard contains optics necessary to implement an I^3 Sensor with up to four subapertures, a 20 Hz to 2.5 kHz nutator, visible and IR reference and detector systems for system tests in both regimes, a source simulator having controllable angular size, brightness and contrast, scanner mirrors for injection and correction of calibrated source tilt variations, various accessory drivers and modulation systems, and a four-channel electronic version of the I^3 Sensor processor. The entire

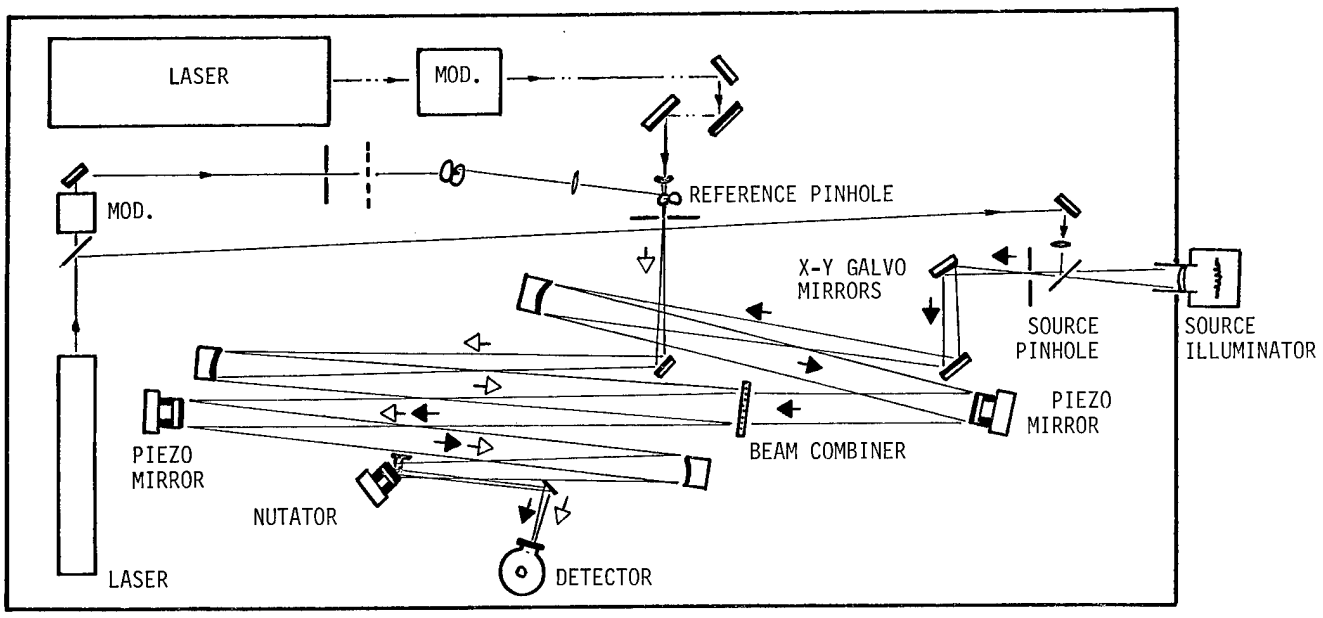


Figure 2. I^3 Sensor optical breadboard. The optical components for the I^3 Sensor and associated test assemblies are shown approximately to scale, as laid out on a 4'x 8' optical table. The two important optical paths are the source (dark arrow) and reference (light arrow) paths. A source mask containing a replaceable pinhole is illuminated, and the resulting beam collimated to provide a test wavefront. Time varying tilt can be added using the x-y galvos and then removed using the piezo mirror. A similar reference pinhole is laser illuminated and its beam collimated. The reference beam is added to the source beam and the pair are focussed, nutated and detected.

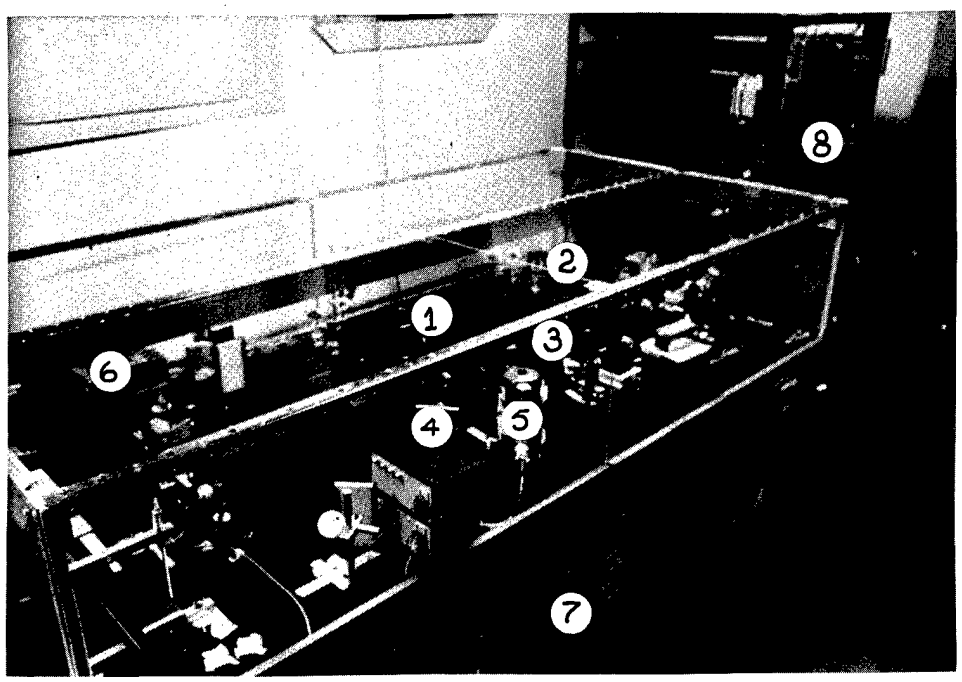


Figure 3. The I^3 breadboard and associated equipment. Numbers indicate locations of: 1) reference pinhole, 2) source pinhole, 3) beam combiner, 4) nutator mirror, 5) IR detector, 6) reference laser, 7) HV amplifiers, and 8) I^3 Processor. Plastic cover reduces ambient turbulence and contains stray IR beams.

system can be monitored and controlled by a Data General Nova 3/12 minicomputer equipped with a DG/DAC digital-analog interface[4]. Figure 3 shows the breadboard and some peripheral equipment.

Experimental results: sensitivity

One of the primary requirements for a wavefront sensor in phase conjugate correction and tracking applications is system sensitivity. For any system based on a spot position sensing or analogous technique (which includes pure Hartmann systems, lateral shearing interferometers and the I^3 Sensor) the ultimate phase measurement accuracy is related to the signal-to-noise ratio (SNR) during the observation period by:

$$\sigma = \frac{1}{SNR}$$

where σ is the variance in waves in the phase measurement.

Extensive testing has been carried out to determine how well the I^3 Sensor can, in practice, approach this limit. These results are summarized in Figure 4. The open points were obtained using the electronic emulator to drive the I^3 Sensor processor; with this system and the computer it is quite easy to develop statistical data over a wide range of signal-to-noise ratios. The closed points represent data taken using the full I^3 Sensor breadboard and processor. All data were taken with optimal nutation radius.

The close agreement between theoretically predicted (solid line) and experimentally obtained phase accuracies show that the I^3 Sensor can give optimal low light level performance in a buildable system.

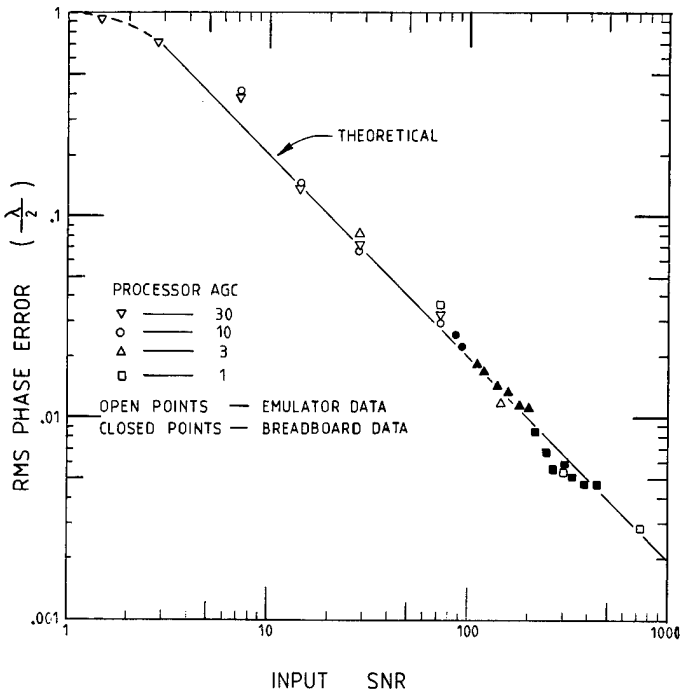


Figure 4. Signal-to-noise performance of the I^3 Sensor. Residual rms phase error in half-waves is given as a function of input SNR over a 1 Hz bandwidth. The open points were obtained using the electronic quad cell emulator and the analog I^3 processor; the closed points were taken with the analog processor and the breadboard. Indicated input amplifier gains are normally set by the system AGC. The solid line is the expected inverse relationship to SNR (see text) with the rollover at SNR = 1 an observed variance limit, since phase errors greater than $\lambda/2$ alias to within that value.

Closed-loop experiments

In order to verify the I^3 Sensor system parameters and their interactions, a series of closed-loop tilt compensation experiments was carried out. Three implementations of the

I^3 processor have been used:

1) an all electronic system using switched analog integrators for the primary data extraction steps and analog computation circuitry for the arithmetic section,

2) a hybrid configuration[4] which performed the integrations in the analog hardware but did the arithmetic processing in the minicomputer, and

3) an all-software processing system[4] in which the detector waveforms were amplified and fed directly to the computer, which generated the error signals for applications to the compensator mirror.

Each version offers a different combination of speed and flexibility in testing modifications of the basic processing algorithm. All of them can be monitored by the computer for real-time data collection and performance analysis.

Figure 5 shows the results of one experiment using the hybrid processing system to correct a 1 Hz sinusoidal tilt jitter. The tilt variation was injected into the source beam x-axis by a galvanometer mirror, and correction was applied to both axes of the piezoelectric mirror (see Figure 2). A nutation rate of 483 Hz was set by an integrator implemented in software to form a Type 1 servo system with a 75 Hz unity gain frequency. Since a Type 1 servo has theoretically infinite loop gain at zero frequency, the observed static offset is completely removed, and the sinusoidal jitter almost so (loop gain at 1Hz is 75).

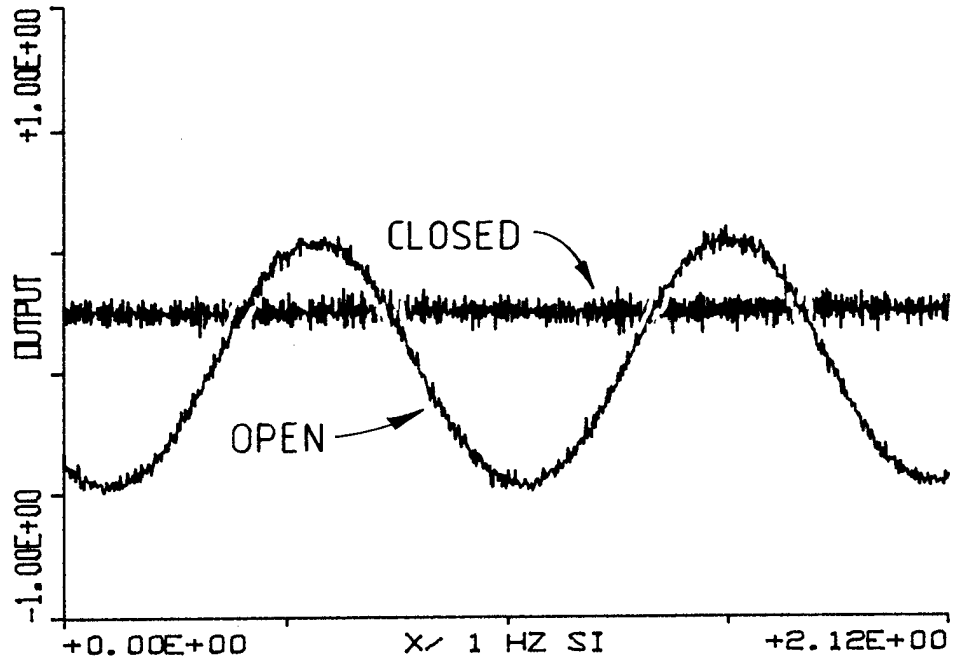


Figure 5. Open and closed-loop jitter correction results for the hybrid I^3 Sensor the open-loop phase input shows a 1 Hz sinusoid with static offset, which is almost completely corrected on loop closure (residue can only be detected in FFT). The vertical scale is approximately ±1 wave of tilt, while the horizontal (time) scale represents a sample duration of 2.12 seconds.

The same set-up was also used to correct local air turbulence which was generated by appropriate placement of a fan and heater near the source path. Figure 6 shows the measured phase output variation due to the turbulence over a period of about two seconds; very good correction results when the loop is closed. The reason for this can be seen in Figure 7, which gives the spectral power density of the measured phase disturbance in the region below 60 Hz. An approximately 1/f distribution on the phase noise is seen (an exact -5/3 power dependence, as in a Kolmogorov distribution, is not necessarily expected from this experimentally contrived turbulence source) in the open-loop measurement. This is ideally matched by the 1/f response of the integrating filter, yielding a nearly spectrally flat residual noise level.

Another feature important to closed-loop use is the I^3 Sensor's unique response to image contrast. The nutation modulation technique provides an AC carrier which uniquely encodes

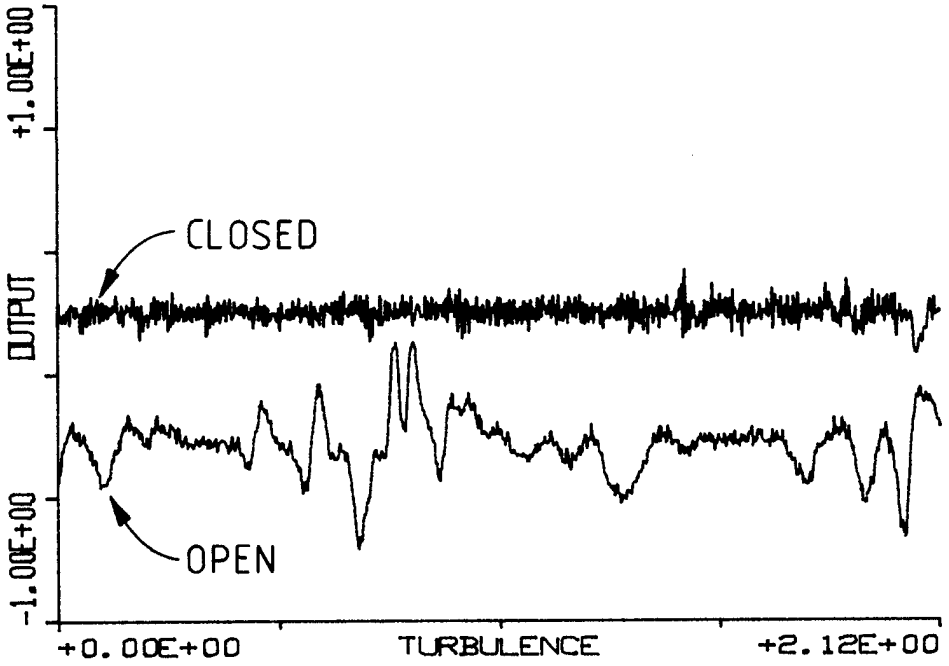


Figure 6. Closed-loop correction of air turbulence. The open-loop trace shows phase fluctuations due to heated air crossing the optical path; the offset is a true static tilt. No direct relationship should be inferred from the events seen in each trace, since the data were taken on successive runs, so that the noise is uncorrelated.

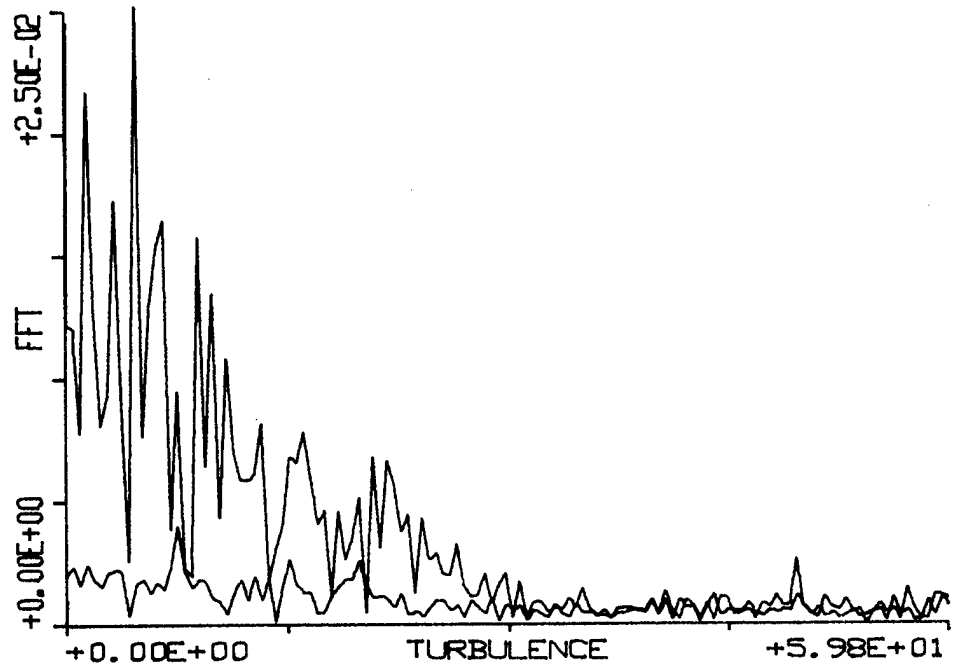


Figure 7. Power spectra of the turbulence data in Figure 6. The upper trace shows the uncorrected phase noise density due to the turbulence (the zero frequency peak is truncated to fit the plot). The lower closed-loop spectrum shows the near optimal correction of the "1/f"-like noise by the Type 1 digital servo. Both data sets are plotted with the same arbitrary power units; the horizontal scale is frequency in Hz.

positive and negative image power. That is, the source wavefront can originate from a bright feature in a dark background (positive contrast) or a dark feature on a bright background. By Babinet's principle, the same information should be present, and in fact the I^3 Sensor supplies correct wavefront information in either case with no operator intervention required as the contrast changes sign.

This has been experimentally verified in this way: a source is modeled by a room temperature plate containing a small pinhole, and the pinhole is illuminated by radiation from a hot body source. A phase jitter correction loop is then closed, using this source. The hot body is then replaced by an object cooled to liquid nitrogen temperature so that less radiation emerges from the pinhole than from the surrounding area, yielding a negative contrast source. With no adjustment of system at all, closed-loop operation is immediately regained; that is, the sign of the phase error signal from the I^3 Sensor is correct, regardless of the sign of the source intensity. The signal-to-noise performance of the system is governed by the curve of Figure 4 in each case.

We have presented some of the results obtained during the development and evaluation of the I^3 wavefront sensor. Emphasis has been placed on experimentally verified characteristics which have direct impact on the optical feedback systems of which the I^3 Sensor would be a component. Although some of the work is currently being extended and a number of special application studies are in progress, the understanding of the device parameters required for system use has been achieved.

Acknowledgment

This paper is based on work sponsored by: The Air Force Weapons Laboratory (AFWL), Air Force Systems Command, United States Air Force, Kirtland Air Force Base, New Mexico 87117.

References

1. U. S. Patent No. 4,141,652.
2. L. E. Schmutz, J. K. Bowker, Julius Feinleib, S. Tubbs, "Integrated Imaging Irradiance (I^3) Sensor: A New Method for Real-Time Wavefront Mensuration," Proc. of the SPIE, 179 (1979).
3. J. K. Bowker, J. Feinleib, S. Landon, L. E. Schmutz, S. J. Tubbs, "I^3 Sensor Wavefront Performance," Proceedings of SPIE, 190 (1979).
4. Susan N. Landon, L. E. Schmutz, Steven J. Tubbs, "Real-time minicomputer control of the I^3 wavefront sensor," SPIE Tech. Symposium East, Active Opt. Devices and Appl., (1980).

The effect of wavefront sensor characteristics and spatiotemporal coupling on the correcting capability of a deformable mirror

Robert K. Tyson, Dale M. Byrne
United Technologies Research Center, Optics and Applied Technology Laboratory
P.O. Box 2691, West Palm Beach, Florida 33402

Abstract

By modeling a wavefront sensor and deformable mirror system as a high pass spatial and temporal frequency filter, its correcting ability in terms of a transfer function can be determined. Since the characteristics of a particular sensor/mirror system, as well as the uncorrected incident wavefront, affect the form of the transfer function, both must be used to describe the system's correcting ability. In this paper a mathematical formalism is presented that is based upon a linear systems approach and which relates the transfer function to the characteristics of a deformable mirror system. Three methods of introducing system characteristics into the filtering process to correct for random and deterministic errors are presented. In situations where spatiotemporal coupling becomes important, a three dimensional filter in frequency space is employed.

Introduction

In recent years there has been considerable effort expended in the development of deformable mirrors for use in active optical systems to perform real time correction of wavefront errors. Feinleib, et al.[1] have reported the development of a Monolithic Piezoelectric Mirror (MPM) which was used in a prototype system to correct atmospherically degraded wavefronts. Angelbeck, et al.[2] have developed a deformable mirror system to be used as a phase correction device for a higher power cw laser operating at 10.6 μm. This system has 52 discrete actuators which perturb a thin molybdenum surface. Pearson and Hansen[3] have presented experimental results of two different deformable mirror designs, both of which used discrete actuators to produce local surface deformations in a thin front surface mirror. Specific attention was given to the temporal response of the mirrors as well as the wavefront correcting capability when used in a multidither COAT* system.

Attention has also been directed toward the analysis of deformable mirror systems as to their ability to correct wavefront errors. O'Meara[4] has presented a theory that describes the operation of a multidither adaptive optical system which uses a deformable mirror as a wavefront corrector. Both the spatial and temporal aspects of the system, as well as resulting phase ambiguities, are considered in this analysis. Wagner[5] considered the deformable mirror wavefront correction process as a spatial filtering operation. The deformable mirror is capable of modifying its surface profile to match that of a distorted wavefront only for wavefronts having spatial frequencies that are less than a critical value. This critical value is primarily a function of the actuator spacing. Harvey and Callahan[6] have extended the filtering concept of deformable mirror operation and investigated the manner in which the actuator density affects the residual wavefront variance when attempting to correct for certain deterministic wavefront errors. A Hartmann based wavefront sensor and a deformable mirror having a rectangular array of actuators were modeled and were used to compute the deformable mirror transfer function which filters the frequency spectrum of the incident wavefront.

In this paper the spatial frequency filtering operation of a deformable mirror will be extended further to include characteristics of the deformable mirror (interactuator spacing and actuator influence functions). A mathematical formalism which describes the filtering operation in terms of linear systems theory will be developed. To account for both spatial and temporal fluctuations of wavefront errors, a power spectral density approach will be described and its application will be illustrated. Approximations for the transfer function will be presented that allow the method to be used to predict the corrective capability of realistic deformable mirror systems.

The Transfer Function

As stated previously both Wagner[5] and Harvey and Callahan[6] have treated the wavefront correction process of a deformable mirror as a spatial filtering operation. It can be readily seen that for a rectangular array of actuators, the highest frequency obtainable by moving alternate actuators, with spacing a, in opposite directions is given by $f_c = 1/2a$. Therefore for incident wavefronts having spatial frequencies less than f_c, reasonable wavefront correction would be expected. However, for incident wavefronts having spatial frequencies higher than f_c, no correction would be possible. Hence, the deformable mirror acts to filter out those frequencies less than f_c while passing unmodified those greater than f_c. The idealized filter described above is shown in Fig. 1.

Mathematically the filtering operation can be described by the expression:

$$\tilde{W}_R(f_x,f_y) = \tilde{H}(f_x,f_y)\ \tilde{W}(f_x,\ f_y), \tag{1}$$

* Coherent Optical Adaptive Technique

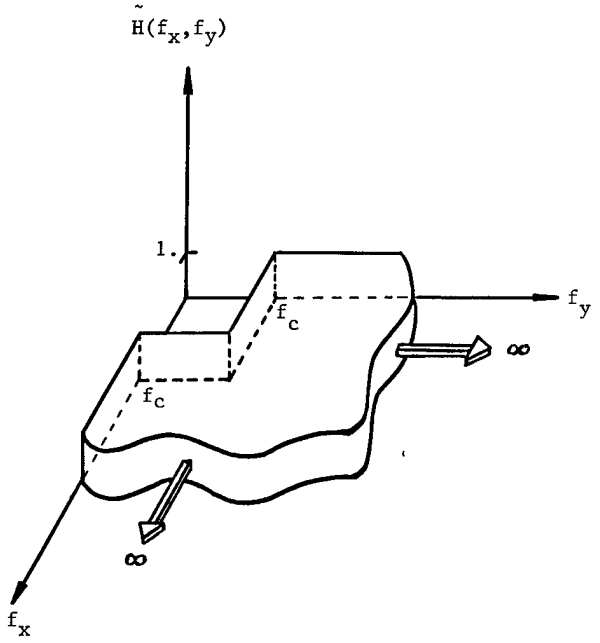


Figure 1. A two dimensional idealized transfer function having a cut-on frequency of f_c.

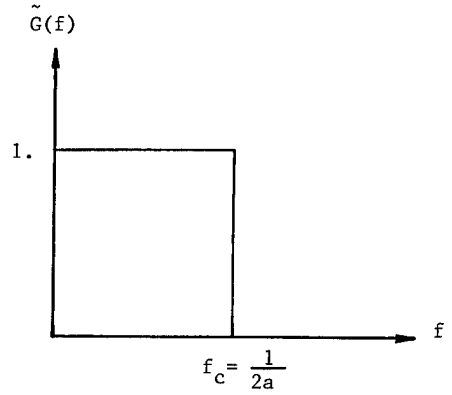


Figure 2. Auxiliary function, $\tilde{G}(f)$, for the idealized transfer function.

where f_x and f_y are spatial frequencies in the x and y directions respectively $\tilde{W}(f_x,f_y)$ and $\tilde{W}_R(f_x,f_y)$ are the frequency spectra of the incident and corrected (or residual) wavefronts, and $\tilde{H}(f_x,f_y)$ is the filter or transfer function. The parallel relationship in the spatial domain is

$$W_R(x,y) = h(x,y)**W(x,y), \tag{2}$$

where $W(x,y)$ and $W_R(x,y)$ are the spatial representations of the wavefront errors before and after correction and $h(x,y)$ corresponds to an effective "impulse response" of the deformable mirror system. The double asterisk denotes convolution in the two spatial coordinates.

Instead of working directly with the transfer function, an auxiliary function $\tilde{G}(f_x,f_y)$ will be defined as

$$\tilde{G}(f_x,f_y) = 1 - \tilde{H}(f_x,,f_y). \tag{3}$$

Therefore, the impulse response becomes

$$h(x,y) = \delta(x)\ \delta(y) - g(x,y), \tag{4}$$

where $g(x,y)$ is the inverse Fourier transform of $\tilde{G}(f_x,f_y)$. The expression for the residual wavefront error which exists after correction now becomes

$$W_R(x,y)=h(x,y)**W(x,y) = W(x,y) - g(x,y)**W(x,y). \tag{5}$$

When written in this form it is easy to see that the convolution of $W(x,y)$, with $g(x,y)$ is the mathematical representation of the modification produced by the deformable mirror to the incident wavefront. The corresponding expression in the spatial frequency domain is

$$\tilde{W}_R(f_x,f_y) = \tilde{W}(f_x,f_y) - \tilde{G}(f_x,f_y)\ \tilde{W}(f_x,f_y). \tag{6}$$

A one dimensional profile of $\tilde{G}(f_x,f_y)$ is shown in Figure 2 for the idealized filter described previously. Again, the filtering effect of the deformable mirror becomes apparent by the removal of all spatial frequencies less than f_c.

Deformable Mirror Characteristics

Consider a deformable mirror with a system of N actuators. When the k^{th} actuator is moved, it produces a perturbation of the mirror surface described by the "influence function" $I_k(x,y)$. It will be assumed in this analysis that actuator movements are sufficiently small so that when more than one actuator is moved, the

resultant mirror surface displacement is a linear superposition of the surface perturbations produced by each actuator. The resultant mirror surface function can therefore be expressed by

$$M(x,y,t) = \sum_{i=1}^{N} A_i(t)\ I_i(x,y), \tag{7}$$

where the $A_i(t)$'s are the displacements of the respective actuator movements. The displacements are shown to be functions of time to account for temporal fluctuations. The influence functions are normalized to a peak value of unity. A wavefront that is initially described by $W(x,y)$ will, after reflection from the deformable mirror, be described by $W_R(x,y)$, which can be written as

$$W_R(x,y) = W(x,y) - 2\cos\Theta\ M(x,y), \tag{8}$$

where Θ is the wavefront angle of incidence. The time dependence has been omitted for simplicity. Equation (8) can be rewritten in terms of the actuator incluence functions as

$$W_R(x,y) = W(x,y) - 2\cos\Theta \sum_{i=1}^{N} A_i I_i(x,y). \tag{9}$$

By comparison of equations (5) and (9), it is readily apparent that the convolution operation in equation (5) is merely the linear superposition of actuator influence functions, i.e.,

$$g(x,y)**W(x,y) = 2\cos\Theta \sum_{i=1}^{N} A_i I_i(x,y). \tag{10}$$

By taking the Fourier transform of equation (10) and dividing by $W(f_x,f_y)$, an expression for the auxiliary function $G(f_x,f_y)$ can be obtained:

$$\tilde{G}(f_x,f_y) = 2\cos\Theta \frac{\sum_{i=1}^{N} A_i \tilde{I}_i(f_x,f_y)}{\tilde{W}(f_x,f_y)}. \tag{11}$$

where $\tilde{I}_i(f_x,f_y)$ is the Fourier spectrum of the i^{th} actuator influence function. By referring to equation (3), the transfer function of the deformable mirror system is formally determined by the expression

$$\tilde{H}(f_x,f_y) = 1-2\cos\Theta \frac{\sum_{i=1}^{N} A_i \tilde{I}_i(f_x,f_y)}{\tilde{W}(f_x,f_y)}. \tag{12}$$

Inspection of equation (12) reveals that the transfer function depends upon the Fourier spectrum of all the actuator influence functions, not just the inter-actuator spacing. Hence the idealized form of a cut-on transfer function shown in Figure 1 should not be expected to apply to a realistic deformable mirror. The transfer function is also dependent upon the spectrum of the incident wavefront as well as the displacements of the individual actuators. Ideally, if the mirror surface could be made to conform to the shape of the incident wavefront (neglecting a factor of 2 and assuming near normal incidence), then the ratio on the right side of equation (12), $G(f_x,f_y)$, would be unity and the transfer function would be identically zero. This is tantamount to $g(x,y)$ being a delta function in the spatial domain resulting in perfect wavefront correction.

In the classical linear systems theory, the transfer function depends only upon the system characteristics and not upon the input. However for adaptive systems, the response of the system must be dependent upon the input function (in this case wavefront error) for the system to operate optimally. In the case of a deformable mirror correction system, if the Fourier spectrum of the incident wavefront error were known and the actuator displacements were determined, then the performance of the deformable mirror system could be described in terms of classical linear systems concepts. A transfer function can be defined that multiplies the Fourier spectrum of the incident wavefront to produce the spectrum of the corrected wavefront. The application of these concepts will be described in subsequent sections.

Sensor/Deformable Mirror Characteristics

The incorporation of realistic sensor or deformable mirror characteristics into the linear filter method requires either direct calculation of the filter function from known experimental data or approximations of the filter function to simulate the sensor and/or deformable mirror. The exact calculation of equation (12) requires knowledge of the actuator amplitudes, their influence functions, and the uncorrected wavefront. As stated above, this method is designed for a system analysis without the requirement that all possible uncorrected (incident) wavefronts be known. That being the case we have taken the approach that system characteristics can be approximated by a perturbation on an ideal, step filter function. The most basic perturbation is simply an additive term applied to the resultant wavefront error, where the magnitude is dependent upon some system characteristic such as signal-to-noise ratio, temporal bandwidth, or structural resonances. An alternative approach requires modification of the filter function by some predetermined amount which can be dependent upon sensor or deformable mirror characteristics. If the ideal filter is a step function with cut-on frequency 1/2a (where a is the actuator spacing), a "less-than-ideal" filter can be described as a step frequency 1/2A (where A, the effective actuator spacing is greater than a). This new "effective filter function" has the same properties as the ideal filter except for the larger passband. The effective actuator

spacing can be found from experiment.

A more general extension of the previous method assumes a filter function that is not a step function, but rather a function that is near zero for small spatial frequencies and and gradually increases to unity for large spatial frequencies. An expression for $\tilde{H}$ in one dimension that possesses this form is

$$\tilde{H}(f) = (1 + B \exp(-bf))^{-1} \quad (13)$$

where B and b are constants and f is the spatial frequency.
The amount of correction for the low frequencies is determined by the coefficient B whereas the exponential term describes the slope of the rise. The equivalent cut-on frequency is found by determining the frequency where $H(f_c) = 0.5$, that is, $f_c = \ln B/b$. The coefficients can then be approximated by determining the lowest order correctability of the adaptive system, calculating B, then relating actuator spacing and B to the coefficient by b=2a ln B. Each approach requires a tradeoff of simplicity vs. accuracy and each should be judged on a case-by-case basis.

Temporal Dependence

Since most adaptive systems are not static correction systems alone, the dynamics of the sensor/deformable mirror subsystem must be incorporated. The classical approach treats the temporal problem unilaterally with the assumption that temporal characteristics are the same at every point on the mirror suface. At most, different aberration modes are treated separately, with some hope that the higher order modes act similarly to the lower order modes. However, the structural dynamics of deformable mirrors and zonal control systems indicate that spatiotemporal coupling can be significant.

To find the residual rms wavefront error, one needs only to integrate the Power Spectral Density (PSD) of the wavefront error.[6] The PSD is found by Fourier transforming the autocovariance function of the wavefront error. The wavefront variance is proportional to the volume under the two-dimensional PSD surface. The PSD can subsequently be filtered to obtain a PSD that corresponds to the corrected wavefront, from which the variance of the corrected wavefront can be calculated. The drawback of the PSD approach manifests itself in the non-reversibility of the autocovariance operation. Only the wavefront variance and not the deterministic corrected surface can be found. The extension of the two spatial dimension PSD to a three dimensional spatiotemporal PSD is straightforward. Along any axis the PSD for that particular frequency variable can be found. The off-axis values in the PSD volume are the cross-correlations between spatial and temporal characteristics. The integral of the non-homogeneous volume is proportional, once again, to the wavefront variance. This process is analagous to calculating the mass of a non-uniform density distribution. The integral of the density function is the desired quantity, in this case, wavefront variance.

If the correction or filtering occurs, a volume is extracted from the low frequency center of the PSD, reducing the total integral by the amount related to the correcting capability of the sensor/deformable mirror system. It should be pointed out that this approach is valid for non-stationary as well as non-ergodic systems, where other approaches may not.

Illustration of a 3-D PSD is difficult since the function $PSD(f_x,f_y,\nu)$ is four-dimensional (ν is the temporal frequency variable). However, if one spatial degree of freedom is suppressed a PSD (f,ν) might appear as in Figure 3.

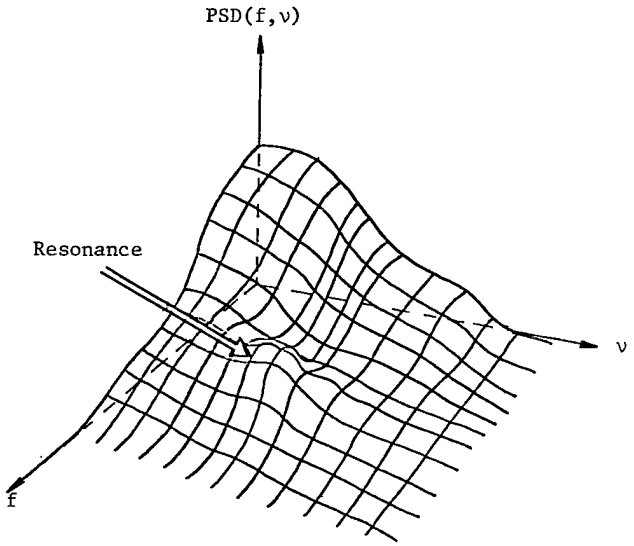


Figure 4. A representative wavefront PSD illustrating spatiotemporal coupling.

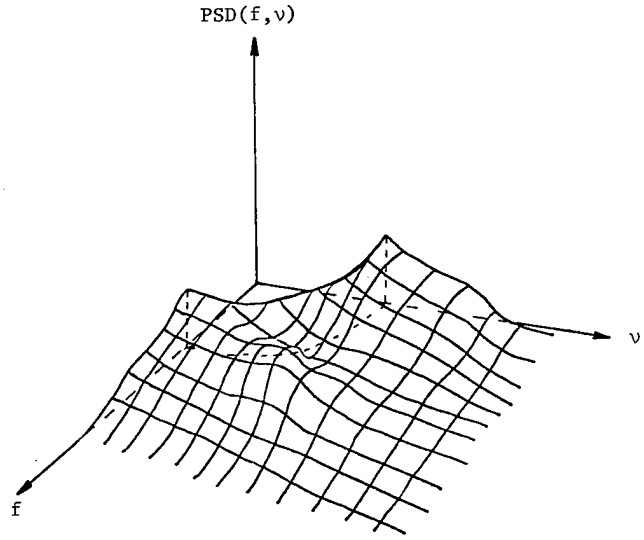


Figure 3. The PSD of Figure 3 after correction by an idealized transfer function.

Notice the resonance appearing in Figure 3 where enhanced spatiotemporal coupling occurs. This represents electromechanical couplings across deformable mirror surfaces when the temporal frequency of the actuators interacts with the structural resonances of the surface. This resonance will not appear by treating each one dimensional PSD separately.

If the PSD in Figure 3 were applied to an adaptive optical system, the corrected PSD might appear as in Figure 4. Clearly the integral under the surface is less than that of the uncorrected PSD, illustrating the correcting capability of the deformable mirror.

Conclusion

Based upon the work of Wagner[5] and Harvey and Callahan[6] in which deformable mirror correcting ability was analyzed in terms of a spatial filtering process, we have developed analytic expressions for the deformable mirror transfer function. The transfer function was shown to depend on the spatial frequency content of the actuator influence functions, the magnitudes of their respective displacements and on the spatial frequency content of the incident wavefront. Approximate methods were presented that allowed the spatial frequency method to be applied in order to predict the performance of a realistic adaptive optical system. Although rather difficult to implement, a three dimensional Power Spectral Density approach to the adaptive compensation process becomes necessary where spatiotemporal coupling causes a breakdown in the assumptions for a stationary, ergodic, or symmetric system.

References

1. Feinleib, J., S. J. Lipson and P. F. Cone, "Monolithic Piezoelectric Mirror for Wavefront Correction," Appl. Phys. Lett., 25, 311, (1974).

2. Angelbeck, A. W., R. C. DePietro, E. D. Hasselmark, E. Gagnon, A. F. Greiner, W. C. McClurg, G. R. Wisner and R. H. Freeman "Development of Cooled Metal Mirror with Active Figure Control, Vols. I and II, " United Technologies Research Center Report R75-911804-24, December, 1975, pp. 400.

3. Pearson, J. E. and S. Hansen, "Experimental Studies of a Deformable Mirror Adaptive Optical System, "Jour. Opt. Soc. Amer., 67, 325, (1977).

4. O'Meara, T. R., "Theory of Multidither Adaptive Optical Systems Operating with Zonal Control of Deformable Mirrors," Jour. Opt. Soc. Amer., 67, 318, (1977).

5. Wagner, R. E., Imagery Utilizing Multiple Focal Planes, Ph.D. Dissertation, University of Arizona, 1976.

6. Harvey, James E. and G. M. Callahan, "Wavefront Error Compensation Capabilities of Multi-Actuator Deformable Mirrors," Proc. S.P.I.E., Adaptive Optical Components, 141, 141-08, S. Holly and L. James, eds., (1978).

Improved figure control with edge application of forces and moments

D. M. Aspinwall, T. J. Karr
Lockheed Missiles & Space Company, Inc.
Palo Alto Research Laboratory, Palo Alto, California 94304

Abstract

A new actuation scheme for mirror figure control, introducing a complete set of membrane shear and bending stresses onto the edges of the mirror, is presented. Thermoelastic response of both segmented and continuous mirrors is analyzed from a closed form solution of the "exact" Love theory of thin spherical shells. There are significant departures from Reissner's theory of shallow shells. Very low residual figure error is possible by edge control alone, lower than with interior point loads. Minimum rms residual error stress profiles are computed. Engineering applications are discussed.

Introduction

Large primary optics are required for a number of advanced missions of current interest. Although large passive telescope mirrors of high figure quality have been made, the figure quality is not adequate for all purposes at short wavelengths, and the mass and fabrication time involved may lead to the selection of a lightweight continuous or segmented adaptive mirror to serve as the primary in large optical systems. This paper addresses the figure control, that is the correction of figure errors, in these large mirrors. The figure errors stem from four primary sources: (1) fabrication errors, of most concern for a segmented faceplate; (2) dimensional instability of a reaction structure supporting the actuators and faceplate; (3) thermal environments produced by high-energy laser applications; and (4) gravity release stresses throughout the mirror faceplate and reaction structure.

Since the bandwidth of the errors is quite low, the most straightforward corrective measure to be employed is placing actuators on the faceplate backside and using an adaptive control system to correct the primary mirror figure to the desired accuracy. Many previous studies have addressed the active figure control of large parabolic or spherical faceplates.[1] These studies have usually been based on the shallow shell theory of Reisner which allows prediction of the surface response to piston forces and moments applied at discrete points in the interior of the shell. Applying the forces and moments at discrete points in the interior leaves the edges of the shell free to adopt non-optimum slopes and deflections, as shown in Figure 1. If forces and moments are applied discretely right at the edge of the shell, some constraint is developed, but wrinkles still develop between the actuators.

Full control of the shell boundaries requires a continuous moment, tangential shear, normal force and transverse shear function; this gives a total of four independent functions which must be specified at the shell boundary. Obviously, better edge control will allow smaller residual errors at the expense of more degrees-of-freedom being controlled. The purpose of this paper is to display the improvement in figure control possible with edge control only, that is, control of the shell boundary conditions but no concentrated interior forces and moments.

One of the motivations for addressing this problem was the desire to demonstrate control of the cap of a spherical shell under a linear thermal strain by application of boundary moments only. A thin sphere with a linear thermal gradient through the surface will experience no surface distortion, because the thermal moments are exactly compensated by elastic moments of the same magnitude at every point on the shell.[2] This condition amounts to producing clamped edge conditions for the optics under study. Because the thermal gradients are a most common source of figure error for HEL optics, a control scheme which could compensate for a uniform thermal gradient by applying uniform moments at the edges of a large optic seems very attractive.

Although the angles involved in a 4-m, f/1.25 spherical shell are less than 12 degrees, which usually justifies the use of shallow shell approximations, the Kelvin functions which are solutions of the shallow shell theory do not allow compensation of a uniform thermal gradient by the use of edge moments alone, as they should. Investigation into the cause of this error in the shell partial differential equations showed that the tangential stretching contribution to the change curvature is critical to correcting thermal deformation by using edge moments alone. The theoretical shell response work presented here is based on the exact non-shallow shell theory involving Legendre functions developed by Berry,[3] which includes this stretching contribution.

Thin elastic shell theory

We are concerned with large optics of several meters aperture and up to tens of cm thick. The thermomechanical deformations suffered by this optic in its operating environment are a few λ – large compared to operating tolerances but very small compared to all dimensions of the optic. Under these conditions, the thermomechanical

response of the optic should be an excellent physical approximation to Love's theory of the thin shell. Then, the thermoeleastic stresses in the medium of the mirror can be completely defined by (a) resultant forces and moments acting at the edges of sector-shaped elements of the shell (Figure 2) and by (b) the average temperature and temperature gradient distribution on a surface of the shell. The geometry of the shell gives six strain-displacement equations. Hooke's law gives six stress-strain relations (for an isotropic medium). Demanding static equilibrium at every point in the medium gives five equations of equilibrium. Thus, the thermoelastic behavior of the mirror is described by a system of 17 coupled linear p.d.e.'s, shown in Figure 3 for a spherical mirror.

The boundary conditions for a solution of the p.d.e.'s are the stresses on the edges of the shell. In general, there are five stresses on an edge: on ϕ-constant edges, there are N_ϕ, Q_ϕ, M_ϕ, $N_{\phi\theta}$, $M_{\phi\theta}$; on θ-constant edges, there are N_θ, Q_θ, M_θ, $N_{\phi\theta}$, $M_{\phi\theta}$. These five stresses are statically equivalent to four:

$$N_\phi \qquad\qquad N_\theta$$

$$M_\phi \qquad\qquad M_\theta$$

$$V_\phi = Q_\phi - \frac{1}{R}\frac{\partial M_{\phi\theta}}{\partial \phi} \qquad\qquad V_\theta = Q_\theta - \frac{1}{R\sin\phi}\frac{\partial M_{\phi\theta}}{\partial \theta} \tag{1}$$

$$T_{\phi\theta} = N_{\phi\theta} + \frac{M_{\phi\theta}}{R}$$

that is, a normal force, normal moment, transverse shear, and tangential shear.

The usual shallow shell approximation for spherical shells[4,5] simplifies this system of equations by (a) the geometrical approximation $\sin\varphi \approx \phi$, and by (b) the dynamical approximation of neglecting stretching terms in the change of curvature:

$$\beta_\phi = \frac{1}{R}\left(u_\phi - \frac{\partial w}{\partial \theta}\right) \approx -\frac{1}{R}\frac{\partial w}{\partial \phi}$$

$$\beta_\theta = \frac{1}{R}\left(u_\theta - \csc\phi\,\frac{\partial w}{\partial \theta}\right) \approx -\frac{1}{r}\frac{\partial w}{\partial \theta} \quad . \tag{2}$$

Approximation (b) is justified when the distortions have a curvature comparable to the undistorted curvature of the shell. But when the shell is a large mirror and we are concerned with displacements from its true figure of a few micrometers, smoothly distributed over several meters aperture, then u_ϕ, $u_\theta \approx w$ and equation (1) is not justified. The exact theory of the thin shell must be used to analyze the environmentally-induced figure errors of a large mirror at short wavelengths.

The exact theory for a spherical shell was solved by Berry using the method of a stress function.[3,5] In the complementary solutions, the normal displacement w is a Fourier-Legendre series

$$w = \sum_{i,m} C_m^{\mu_i} P_{\mu_i}^m(\cos\phi) \begin{Bmatrix} \cos(m\theta) \\ \sin(m\theta) \end{Bmatrix} \tag{3}$$

with general Legendre functions P_μ^m:

$$\left[\nabla_m^2 + \mu_i(\mu_i + 1)\right] P_{\mu_i}^m = 0 \qquad , \qquad \mu_i \text{ complex}$$

$$\mu_i = -\frac{1}{2} + \left(\frac{1}{4} - r_i\right)^{1/2} \tag{4}$$

$$r_i^3 + 4r_i^2 + \left[(\xi + 1)(1 - \nu^2) + 4\right] r_i + 2(\xi + 1)(1 - \nu^2) = 0$$

$$\xi = \frac{12R^2}{h^2} \qquad , \qquad \nu = \text{Poisson's ratio} \quad .$$

Tangential displacement u_ϕ, u_θ can be calculated in closed form from equation (3). The particular solutions, the effect of external stress and thermal loads, has the same form as equation (3) with spherical harmonics replacing the general Legendre functions. The edge stresses (boundary conditions) compatible with equation (3)

are linear combinations of P_μ^m, $P_\mu^{m'}$, $P_\mu^{m''}$ with the same expansion coefficients C_{in}^μ scaled by the stiffness constant

$$D = \frac{E\,h^3}{12(1-\nu^2)} \quad , \qquad (E = \text{Young's modulus}) \ .$$

Applications of these edge stresses will result in displacement w, Equation (3). The solutions [Equation (3)] have eight arbitrary constants for every m . This is enough freedom to satisfy the boundary conditions on two edges. Thus, Equation (3) is a complete set of solutions for the thermoeleastic response of a spherical shell with with two boundaries, e.g., a spherical mirror with an obscuration. For a segmented mirror, Equation (3) is one-half of all the solutions since the segment will in general have four boundaries and require 16 boundary conditions.

Two parameters determine how well the mirror can be figured by application of edge stresses alone: ξ and D. Points on the mirror far from the edges respond well to the edge stresses when ξ is small. The mirror responds easily to small edge stresses when D is small. For a given f/#, size, and glass, the only free parameter is the mirror thickness h . The relation of h to ξ, D is shown in Figure 4.

Computer model

We have constructed a computer model of the static thermoelastic behavior of a thin shell spherical mirror. The computer model can handle a spherical mirror any shape as long as its boundaries are ϕ- and θ-constant edges.

The model numerically evaluates the particular solutions [Equation (3)] and determines the distortion of the mirror figure due to external stress and temperature loads. It evaluates the complementary solutions [Equation (3)] and adjusts the arbitrary coefficients C_m^μ to give the best correction of the mirror to the desired figure within the constraint of a maximum correction bandwidth m_o . The criterion for best correction is minimum rms residual normal displacement from the desired figure; tangential displacement is neglected in the correction criterion because its contribution to residual OPD is $\approx\beta \cdot w$, and $\beta << 1$. The edge stresses required for the best-fit correction are calculated also.

The numerical method used in the model is similar to the stiffness matrix method in structural dynamics. The OPD error of the distorted mirror is expanded in spherical harmonics with coefficients S_m^ℓ . Then, the solution coefficients C_m^μ in the expansion [Equation (3)] of residual OPD are found by matrix multiplication of S_m^ℓ . Thus, the model directly relates input OPD to residual (corrected) OPD. The edge stresses are determined by another matrix multiplication on C_m^μ .

Application of model

We have run the model for three representative sources of figure error on a large optic in its operating environment:

I. $\delta R = 250\,\mu m$
II. $\delta T = 250^\circ$, $\delta T' = -200^\circ/m$
III. hot spot

Case I is the SOA precision of fabrication metrology for a large segmented optic. Case II is a severe thermal load in a laser environment. Case III is a serious potential error source in laser applications. We have compared the performance of three mirrors under I–III:

A. f/1.25, 4.0-m aperture, 0.25 obscuration, 10-cm-thick honeycomb with 3mm facesheet, segmented (6 petals) ULE.
B. f/1.25, 4.0-m aperture, 0.25 obscuration, 10-cm-thick solid, continuous ULE.
C. f/1.5, 0.5-m aperture, 0.25 obscuration, 5-cm-thick solid, segmented (6 petals) Zerodur.

Mirror A is near the limit of what can be fabricated with current technology. The honeycomb and very thin facesheet make this mirror respond, for purposes of error correction, like a solid ULE mirror 17 cm thick (ξ = 41,522), while the edge stresses required for correction are comparable to a solid ULE mirror only 5.5 cm thick (stiffness $D = 1.007 \times 10^6$ nt-m). Mirror B is a continuous solid mirror compromise to the correction and stiffness characteristics of A (ξ = 120,000, $D = 5.985 \times 10^6$ nt-m). Mirror C is under fabrication now for a laboratory demonstration of segmented optic figure control by edge actuators later this year. The relation between these three mirrors is shown in Figure 5.

For segmented Mirrors A and C we model only a single segment, since each segment can be controlled independently. We arbitrarily limited the bandwidth of the complementary solutions to $m_o = o$ plus a tilt (m = 1) term. For Cases I and II, this forces $N_{\phi\theta} = M_{\phi\theta} = T_{\phi\theta} = 0$ and corresponds to the simplest type of edge control of only three stress degrees-of-freedom on every edge.

Note that for segmented optics the "best fit" correction computed by the model is not really the minimum rms residual OPD, because the solutions [Equation (3)] over which the model optimizes only span one-half the solution space of a four-sided segment. Thus, the results of our model are only upper bounds for the residual OPD of an edge-controlled segmented mirror.

Discussion of results

The results for Cases I and II are summarized in Tables 1 and 2.

Table 1. Case I

Mirror	OPD Input, Error, rms (μm)	OPD After Piston and Tilt, rms (μm)	OPD After Edge Actuation, rms (nm)	Maximum* Edge Force (nt/m)	Maximum* Edge Moment (nt)
				N_ϕ	M_ϕ
A	250	7.45	0.1	-193	-6
				N_ϕ	M_ϕ
B	250	7.01	0.3	-2,070	56
				N_ϕ	M_ϕ
C	250	7.05	0.009	8,780	-95

*These stresses are applied at the inner radius.

Table 2. Case II

Mirror	OPD Input, Error, rms (μm)	OPD After Piston and Tilt, rms (μm)	OPD After Edge Actuation, rms (pm)	Maximum* Edge Force (nt/m)	Maximum* Edge Moment (nt)
				N_ϕ	M_ϕ
A	75	2.2	30	58	1.8
				N_ϕ	M_ϕ
B	75	2.1	90	620	-16.8
				N_ϕ	M_ϕ
C	18	0.53	0.6	-659	7

*These stresses are applied at the inner radius.

A contour plot of the corrected mirror figure for IA is shown in Figure 6; contour plots of all the mirrors for Cases I and II are similar. All three mirrors can be figured to very precise tolerances by edge stresses alone. But the continuous mirror requires large normal forces on its inner edge to achieve a good figure, which implies a large number of powerful, possibly heavy-edge actuators. For the segmented mirrors, some of the bending error can be corrected by tilting the segment, but this degree-of-freedom is not available to the continuous mirror because there is no tilt component of the input error. Large edge stresses are required to correct Mirror B under this constraint. The segmented mirror does not suffer from this problem. Even for Mirror C, which is small and very stiff, the edge force integrated over the 6.54-cm inner edge is only 43 nt for the severe thermal load of Case II.

The residual figure error for Cases I and II is a series of axisymmetric ripples. The rippling is most severe at the inner edge of the mirror and rapidly smooths out moving away from the edge.

To achieve the optimum correction for Cases I and II, uniform forces and moments must be applied at the inner and outer radii of all mirrors, and rapidly varying normal force and moment must be applied along the straight edges of the segments in Mirrors A and C. The bending moment required for Case I, Mirror A, is shown in Figure 7. This is an optimum moment distribution; departures from it due to a finite number of discrete actuators will degrade the mirror performance. For a segmented optic, rigid-body piston and tilt of each segment can reduce the edge stresses required for figure correction; the small stresses only correct for the OPD left after piston and tilt are done. Thus, a segmented optic is less sensitive than a continuous optic to the less-than-optimum edge stresses applied by real actuators.

The results for Case III are summarized in Table 3. A broad hot spot was applied to Mirror B. The temperature contours are shown in Figure 8, and the residual figure error contours after edge actuation in Figure 9. A much sharper hot spot was applied to Mirrors A and C. The temperature and residual error for Mirror A are shown in Figures 10 and 11.

Table 3. Case III

Mirror	T_{max} (deg)	T_{min} (deg)	OPD Input Error, rms (μm)	OPD After Piston and Tilt, rms (μm)	OPD After Edge Actuation, rms (nm)
A	260.0	241.5	72	21	7
B	260.0	260.0	74	0.84	52
C	255.0	241.5	17	2.5	0.8

The hot spot prints through on the figure of the continuous mirror even after correcting edge stresses are applied. The performance of this mirror is rather marginal even though the hot spot is broad and gentle. The below-optimal action of discrete actuators may degrade this much further.

The hot spot on the segmented mirror is largely erased by the applied edge stresses. The segmented mirrors can perform much better than the continuous mirror in a thermal environment which is more demanding. Mirrors A and C take advantage of the independent rigid-body tilt freedom of each segment which helps correct OPD errors that do not look anything like tilt if examined over the whole aperture. Also, the honeycomb structure gives Mirror A the figure characteristics of a shell much stiffer than A really is, thus improving its edge stress response (Figure 4).

On a continuous mirror, the spatial bandwidth of correction by edge stresses is limited to the bandwidth of the error because each Fourier order is independent of all others. On a segmented mirror, each order is coupled to an infinity of other orders, so correction will always improve as the spatial bandwidth increases. The

only practical limit to correction by edge stresses is the number of actuators that can be put on a segment. Recall that the results in Table 3 were obtained at zero bandwidth. Relaxing this constraint would reduce the residual OPD for Mirrors A and C at the expense of more complex (four stress) actuators.

Conclusions

1. The use of edge actuators to control the boundary conditions of primary mirrors shows a marked improvement in the accuracy of figure correction theoretically attainable as compared to the use of push-pull actuators in the interior of the mirror.

2. A hybrid system which uses some edge actuators for boundary condition control and some interior actuators for localized figure correction will probably be more effective than using either system alone.

3. If a segmented primary mirror is employed to allow fabrication of larger apertures, the radius of curvature mismatch between segments can be controlled to ~ 1 nm with the edge actuation system, thereby loosening the fabrication tolerance on radius of curvature.

4. The theoretically optimum smoothly varying forces and moments required along the mirror boundaries will be difficult to realize in actual hardware designs. There will be some deviation from the ideal functional form for any method of discrete actuation. This favors the segmented optic.

5. As regards hot spots of thermal distortion, the segmented-type primary shows a distinct advantage over the continuous primary in being able to localize and restrict the effect of the hot spot because the spatial frequency of correction is limited only by the number of actuators. The continuous mirror, on the other hand, can be controlled with fewer actuators for symmetrical thermal loadings.

6. The added complexity of edge actuation may prove a stumbling block in applications to practical designs; however, the extremely high accuracy shown in this paper makes the concept very attractive for achieving the very tight tolerances required for short wavelength optical systems.

References

1. David Bushnell, AIAA Jour. 17, 71–77 (1979) and references therein.

2. S. P. Timoshenko and J. N Goodier, Theory of Elasticity, New York, McGraw-Hill, 1970, pp. 452–456.

3. J. G. Berry, Proc. 3rd Midwest Conference on Solid Mechanics, Purdue University Experimental Research Series No. 129, 25–44 (1955).

4. E. Reissner, Jour. Math. Phys. 25, 279–300 (1946); corrrections in 27, 240 (1948).

5. Harry Kraus, Thin Elastic Shells, New York, John Wiley & Sons, 1967, Chapter 6

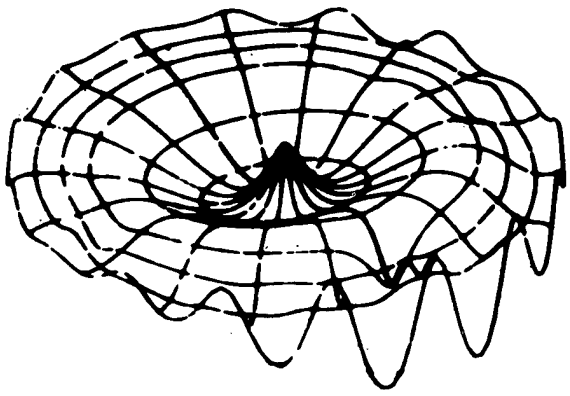


Figure 1. Typical response of a mirror to discrete interior actuators.

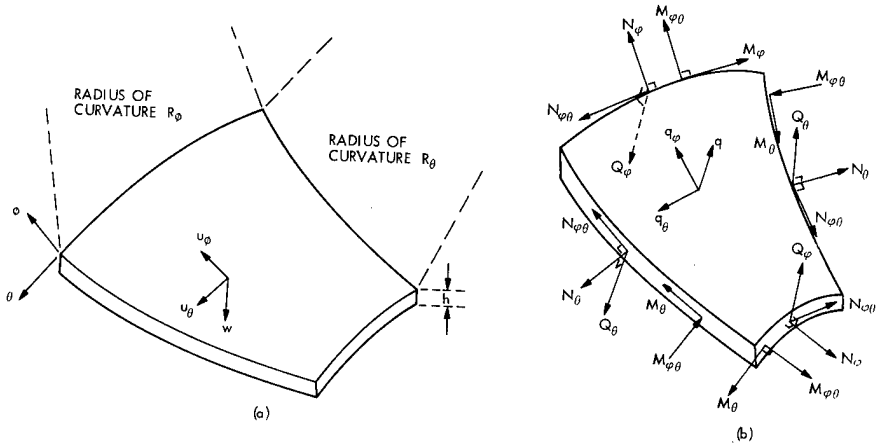


Figure 2. Notation conventions for a thin shell element.
(a) Coordinates and displacements.
(b) Forces and moments.

Strain – Displacement

$$\epsilon^o_\varphi = \frac{1}{R}\left(\frac{\partial u_\varphi}{\partial \varphi} + w\right) \qquad K_\varphi = \frac{1}{R}\frac{\partial \beta_\varphi}{\partial \varphi}$$

$$\epsilon^o_\theta = \frac{1}{R}\left(\csc\varphi \frac{\partial u_\theta}{\partial \theta} + u_\varphi \cot\varphi + w\right) \qquad K_\theta = \frac{1}{R}\left(\csc\varphi \frac{\partial \beta_\theta}{\partial \theta} + \beta_\varphi \cot\varphi\right)$$

$$\gamma^o_{\varphi\theta} = \frac{1}{R}\left(\frac{\partial u_\theta}{\partial \varphi} - u_\theta \cot\varphi + \csc\varphi \frac{\partial u_\varphi}{\partial \theta}\right) \qquad \tau = \frac{1}{R}\left(\frac{\partial \beta_\theta}{\partial \varphi} - \beta_\theta \cot\varphi + \csc\varphi \frac{\partial \beta_\varphi}{\partial \theta}\right)$$

Stress – Strain

$$N_\varphi = \frac{12D}{h^2}\left(\epsilon^o_\theta + \nu\epsilon^o_\theta - (1+\nu)\,\alpha\, T_o\right) \qquad M_\varphi = D\left(K_\varphi + \nu K_\theta - (1+\nu)\,\alpha\, T_1\right)$$

$$N_\theta = \frac{12D}{h^2}\left(\epsilon^o_\theta + \nu\epsilon^o_\varphi - (1+\nu)\,\alpha\, T_o\right) \qquad M_\theta = D\left(K_\theta + \nu K_\varphi - (1+\nu)\,\alpha\, T_1\right)$$

$$N_{\varphi\theta} = \frac{6D(1-\nu)}{h^2}\gamma^o_{\varphi\theta} \qquad M_{\varphi\theta} = \frac{D(1-\nu)}{2}\tau$$

T_o = temperature , T_1 = temperature gradient

Static Equilibrium

$$\frac{\partial}{\partial\varphi}(N_\varphi \sin\varphi) + \frac{\partial N_{\varphi\theta}}{\partial\theta} - N_\theta \cos\varphi + Q_\varphi \sin\varphi + q_\varphi R \sin\varphi = 0$$

$$\frac{\partial}{\partial\varphi}(N_{\varphi\theta} \sin\varphi) + \frac{\partial N_\theta}{\partial\theta} + N_{\varphi\theta}\cos\varphi + Q_\theta \sin\varphi + q_\theta R \sin\varphi = 0$$

$$\frac{\partial}{\partial\varphi}(Q_\varphi \sin\varphi) + \frac{\partial Q_\theta}{\partial\theta} - (N_\varphi + N_\theta)\sin\varphi - q R \sin\varphi = 0$$

$$\frac{\partial}{\partial\varphi}(M_\varphi \sin\varphi) + \frac{\partial M_{\varphi\theta}}{\partial\theta} - M_\theta \cos\varphi - Q_\varphi R \sin\varphi = 0$$

$$\frac{\partial}{\partial\varphi}(M_{\varphi\theta} \sin\varphi) + \frac{\partial M_\theta}{\partial\theta} + M_{\varphi\theta}\cos\varphi - Q_\theta R \sin\varphi = 0$$

q, q_ϕ, q_θ = external forces

Figure 3.

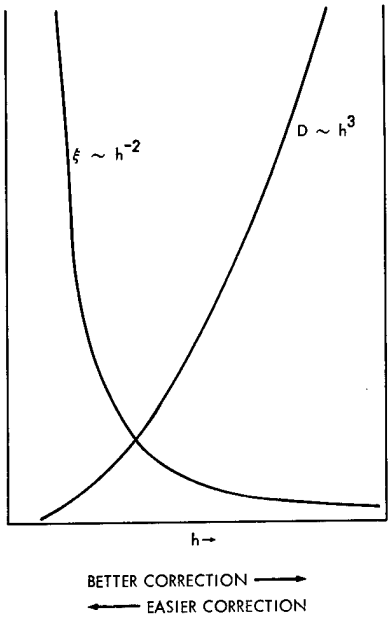


Figure 4.

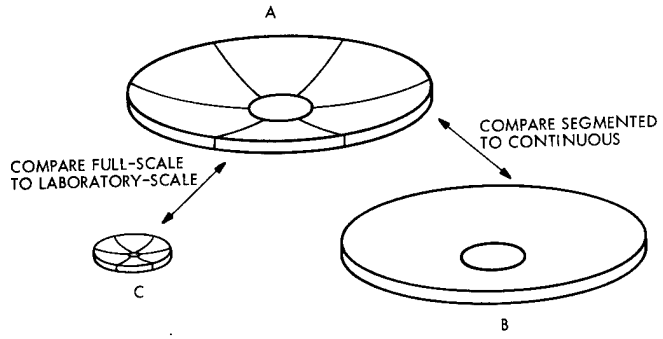


Figure 5.

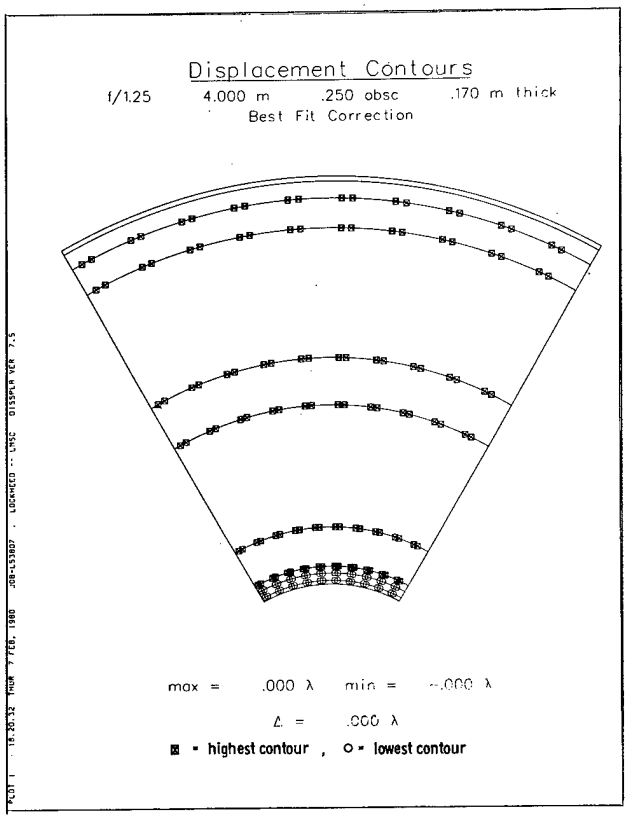


Figure 6.

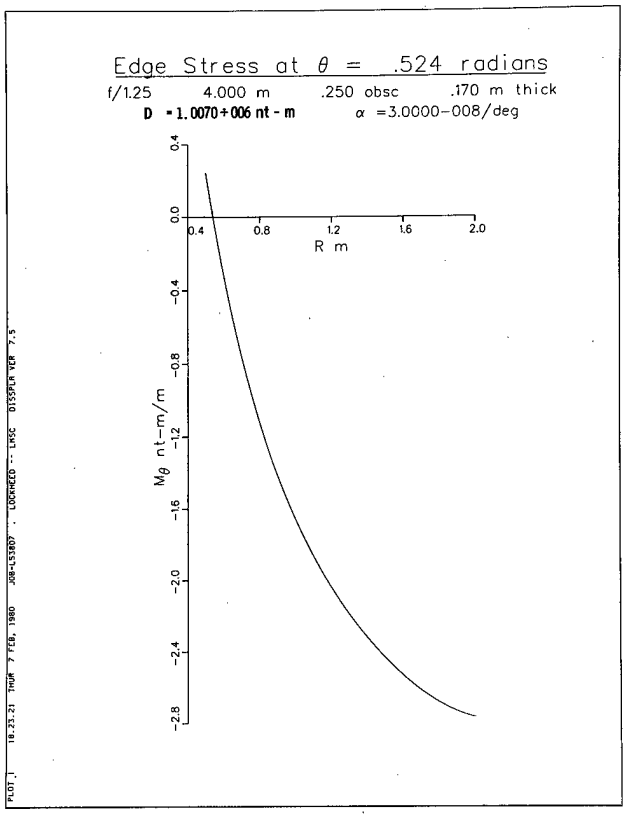


Figure 7.

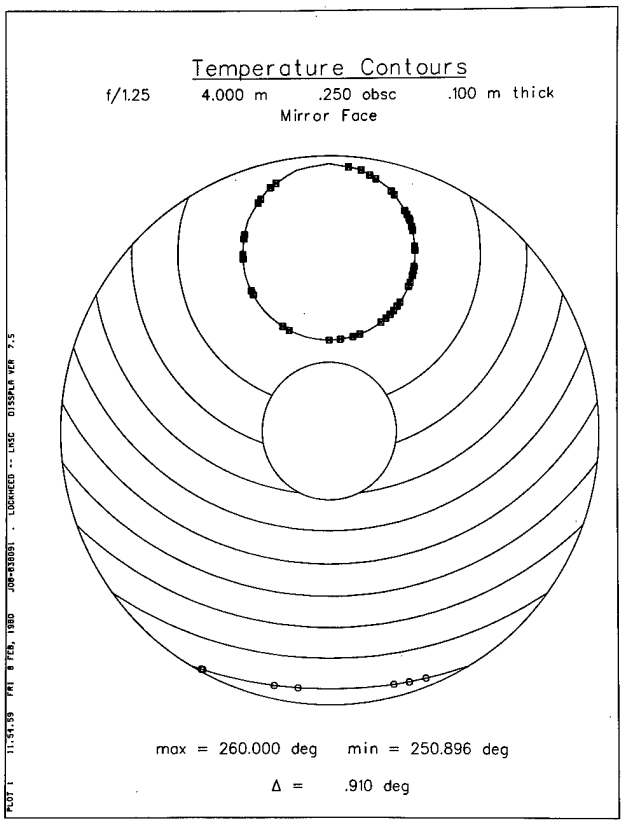


Figure 8.

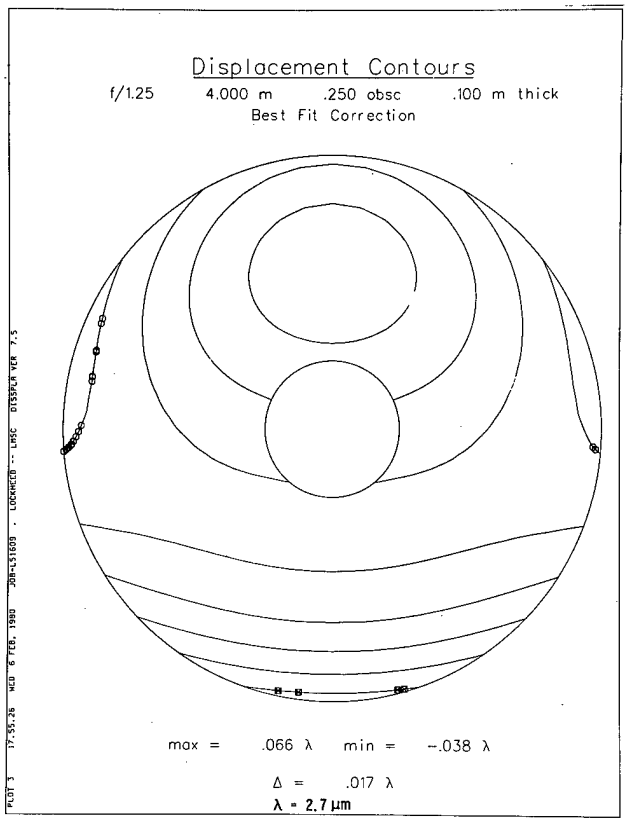


Figure 9.

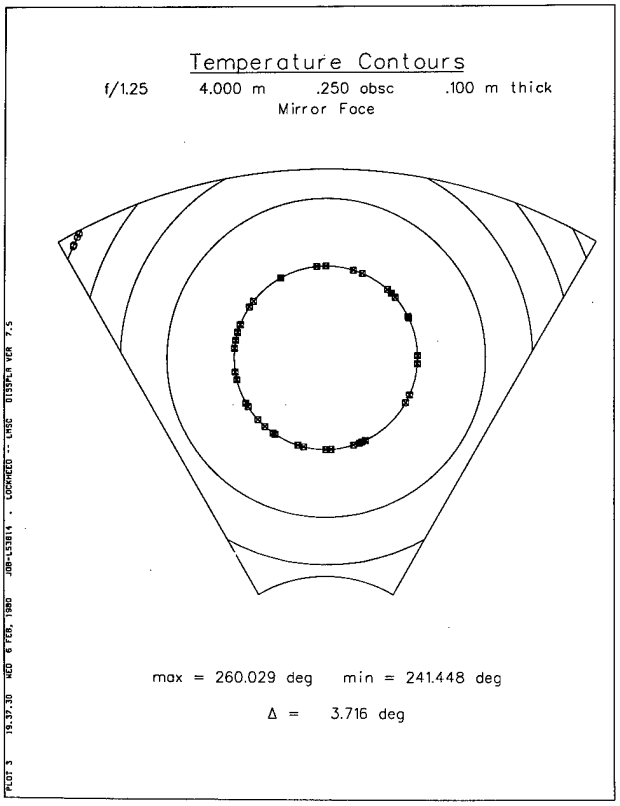


Figure 10.

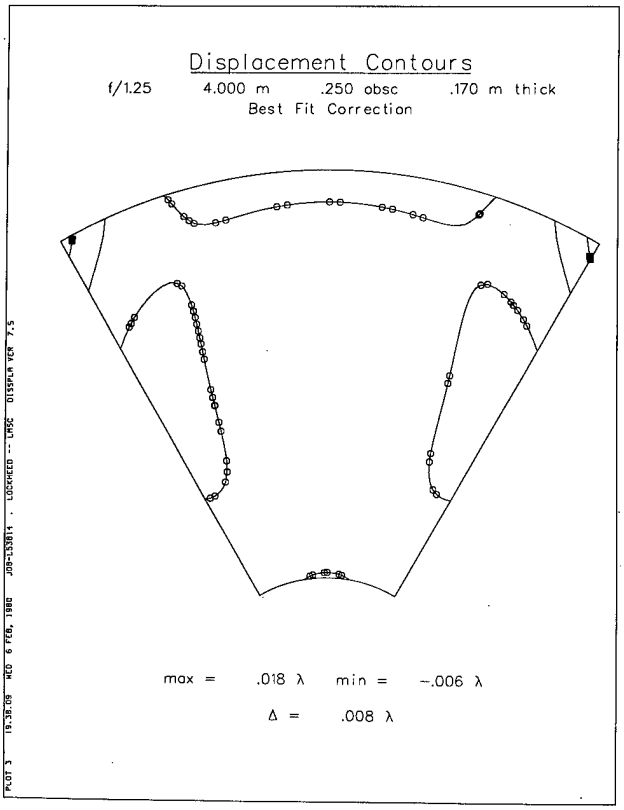


Figure 11.

Device parameters and optical performance of a stacked actuator deformable mirror*

J. H. Everson, R. E. Aldrich, M. Cone, J. Kenemuth****
Itek Corporation, Optical Systems Division, Lexington, Massachusetts 02173

Abstract

A discrete-actuator, deformable mirror has been characterized for use in the infrared spectral region. This device, with a clear aperture diameter of nine inches, achieves a surface deformation of 8.5 μm at 1500 volts. Conformity of the reflective surface with a desired figure has been determined by interferometric analysis and residual error computation for focus, tip, tilt and two orders of astigmatism.

Introduction

With the emergence of larger aperture optical systems requiring diffraction limited performance in the infrared spectral region, new demands have been placed upon deformable surface hardware. To keep pace with requirements of increased surface deformation at reasonable voltages and improved figure stability, we have designed, fabricated and characterized a deformable surface device, which employs discrete actuators made from stacks of piezoelectric (PZT) ceramic. An actuator is energized electrically in parallel and provides series mechanical displacements. The essential features of this mirror are listed below in Table 1.

Table 1. Principal Parameters of the Mirror

Number of Actuators	37	Controlled Aperture	16.5cm
Number of Inert Actuators on Mirror Perimeter	18	Mechanical Resonances	≥300Hz
Clear Aperture	23cm	Average Dilatation	≥8.5μm/1.5kV

Figure 1 presents a conceptual drawing of this device, while Figure 2 displays the actuator array.

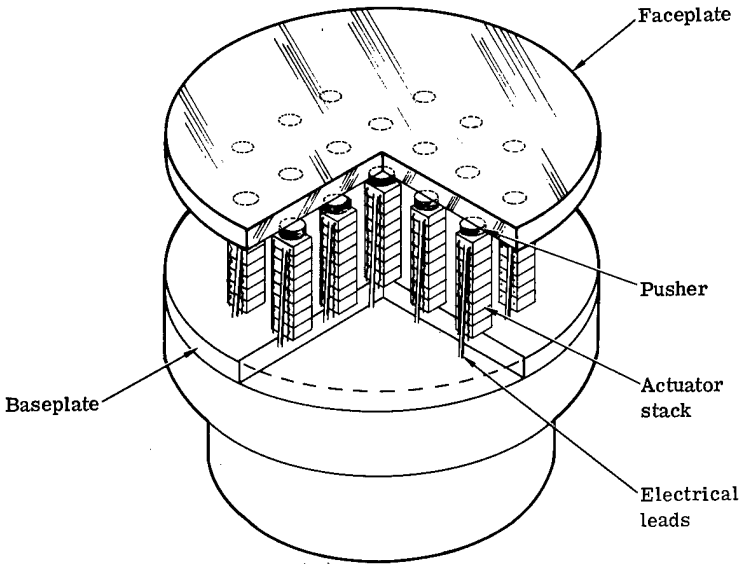


Fig. 1. Concept of stacked actuator deformable mirror.

*This work was sponsored by Contract F29601-78-0060, Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico.

**Dr. J.H. Everson and Dr. R.E. Aldrich are with Itek Optical Systems Division, Maj. M. Cone and Dr. J. Kenemuth are affiliated with Air Force Weapons Laboratory.

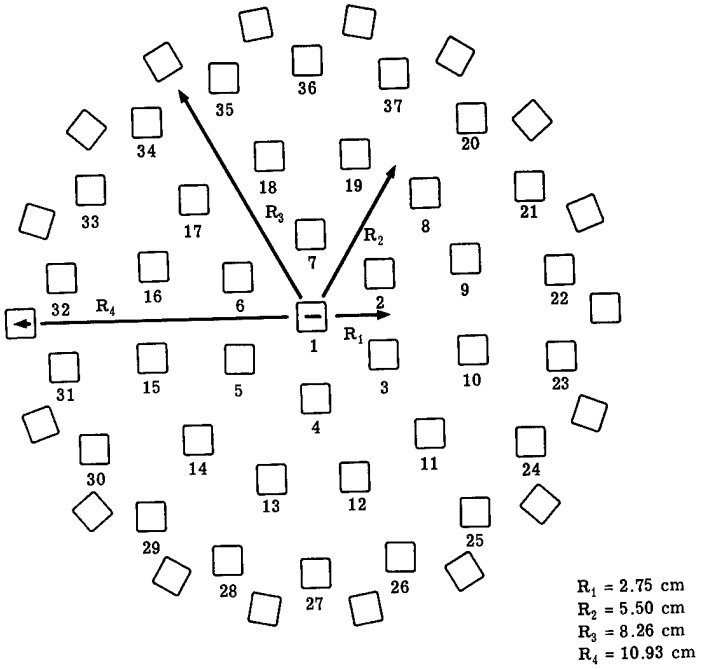


Fig. 2. Numbering scheme for an array of 37 actuators. There are 18 inert actuators bordering the perimeter of the mirror face plate. Drawing is not to scale.

Shown in Figure 3 is the mirror which is seated in a bezel mount and enclosed laterally by a protective dust cover and in the rear by an electrical cover containing a matrix type connector. The mirror itself consists of a glass face plate (which is given a proprietary coating), the actuators and a base plate which is situated in the bezel. Precision machining and special fixturing are utilized to insure coplanar surfaces at the face plate/actuator interface and the base plate/actuator plane at the time of bonding.

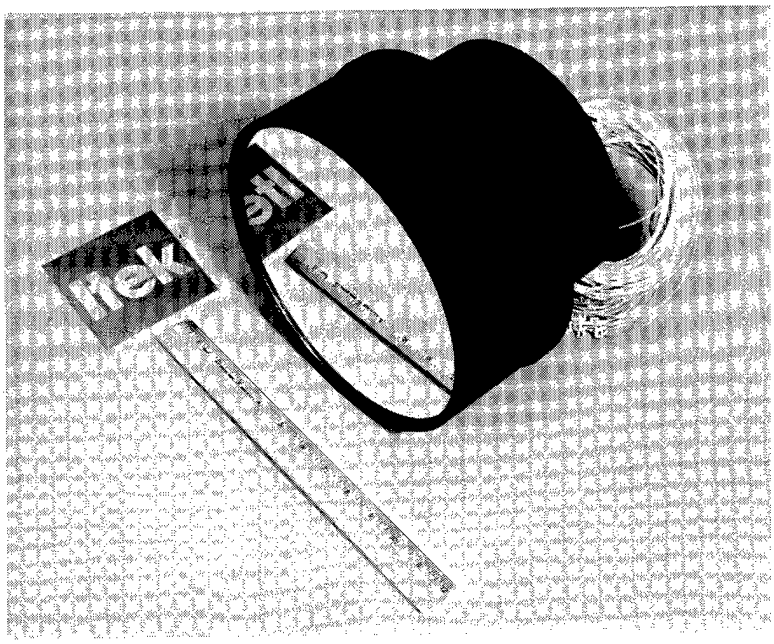


Fig. 3. Deformable mirror in metal housing.

Optical figure stability is optimized by proper actuator design. Unlike some actuator concepts, which operate about a mechanical bias created by spring loading PZT stacks, the actuators for this mirror are epoxy bonded. These bonded actuators offer good, long term stability since they do not experience mechanical creep and depolarization typical of constant spring tension actuators. Thus this mirror does not require individual micrometer adjustments at each actuator to smooth figure degradation. Furthermore, the optical figure is unaffected by orientation of the mirror.

Additional device stabilization has been achieved in terms of ambient temperature changes. Thermal expansion coefficients of materials at critical interfaces have been optimized together with the use of suitable epoxies and bonding techniques.

The sections below discuss device parameters such as surface deformation (influence function), dilatational-voltage response and athermalization measurements. Also presented are performance data which measures the accuracy with which the mirror assumes the shape of low order Zernike polynomials such as focus, astigmatism, etc.

Influence Function

Figure 4 presents a typical influence function from an actuator. The influence function is the surface deformation measured radially from the center of an energized actuator. Surface dilatation is measured by means of a surface profilometer (Dektak), which is capable of measurement to within 50Å or 1 percent. The data of Figure 4 is obtained by scanning at discrete intervals, starting with the energized actuator (#1) and continuing the scan in the direction indicated by the arrow relative to the (unenergized) actuator positions. All unenergized actuators are grounded. Between actuators 6 and 33, the influence function is slightly negative (approximately -4%) and approaches zero at actuator 33. This type of influence function, which is nearly gaussian, creates little "cross talk" at neighboring actuators and therefore simplifies wavefront correction algorithms. The shape and magnitude (sensitivity) of the deformation profile depend upon the actuator force, face plate thickness, interactuator spacing and pusher geometry. The interrelations are complex and at present must be determined experimentally for the specific design.

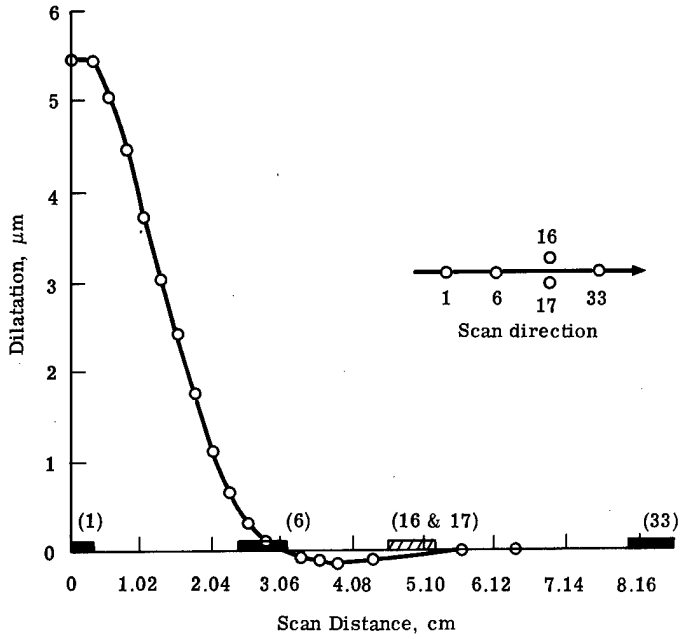


Fig. 4. Influence function of actuator #1 with an applied voltage of 1kV. Glass thickness of mirror face plate is 0.30cm. Scan direction is indicated by the above arrow and actuator numbers.

Dilatation Versus Voltage

The mechanical response of the mirror surface was measured at the center of an actuator for several voltages. Results of these measurements are provided in Figure 5. The data shows a departure from linearity at low voltages (either positive or negative) and at the high voltage range for negative voltages (<-1kV). At ±1.5kV, the deformation is greater than 8 μm for a face plate thickness of 0.30cm. For a final mirror surface thickness of 0.25cm, the average sensitivity is greater than ± 8.5 μm at +1.5kV.

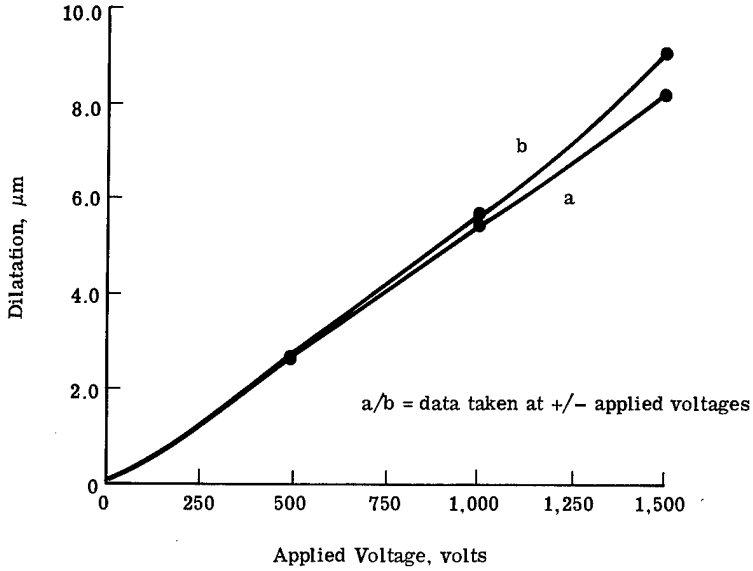


Fig. 5. Typical dilatation versus applied voltage curves for actuator #1. Glass thickness is 0.30cm.

Optical Figure

The optical figure of the mirror was measured at critical times during device fabrication. These times were (1) just after final mirror polish and (2) just after potting the mirror into the bezel. Table 2 shows the final optical figure characteristics for a controlled aperture of 16.5cm.

Table 2. Surface error of mirror. Error is scaled to 3.8 μm.

Peak-Peak	Power	Astigmatism
λ/17	λ/32	λ/38

Temperature Effects

The completely assembeled mirror was subjected to a slowly changing temperature environment to determine what changes, if any, would occur in the optical figure. After installing the mirror in the interferometer (Fizeau type made by Davidson), the temperature of the room was raised to about 27°C and allowed to cool. During the cooling process, Polaroid pictures were taken of the interferograms and the temperature noted at each picture.

From the data the amount of power arising from a temperature difference between 25°C and 22°C is λ/15 at 3.8 μm (wavefront). In other words, there is an incremental contribution to the power at the rate of λ/45/°C, again in terms of wavefront and at 3.8 μm. The validity of assuming the above rate constant to be indicative of the mirror surface behavior has been established by measuring other similar types of deformable mirrors and noting the near linear response over a temperature range of at least 10°C.

Mirror Performance Evaluation

This mirror was then evaluated for its ability to conform to low order Zernike mode shapes. The voltages to be applied to the individual actuators were determined using the displacement required at each actuator location for the desired mode shape. Due to the high uniformity among the actuators and the low coupling between them, the actuators were driven directly to the desired figure. In previous evaluations[1] the voltage to be applied required gain calibration for variations among actuator locations. While it is felt the mirror surface conformity to the desired mode can be improved by iterating on the imposed voltages, no attempt has been made here to do that.

The ability of the mirror to assume the desired Zernike mode shapes was measured by placing the deformable mirror in a Twyman-Green interferometer which had been modified to include an expanding telescope in the sample arm in order to fully illuminate the polished area of the deformable mirror. Interferograms were obtained using 0.6328 micrometer illumination. These interferograms were digitized and then analyzed by fitting the data to a Zernike expansion using a least squares technique. The quantitive figure-of-merit for the surface fidelity was taken to be the RMS deviation of the mirror surface from the desired mode shape.

A collection of interferograms representative of the tilt, focus, and astigmatism modes achieved are shown in Figure 6. The actuator locations are denoted by solid dots. The nominally flat initial condition is shown in the upper left. Only the surface within the active area of the mirror is of significant interest for these evaluations.

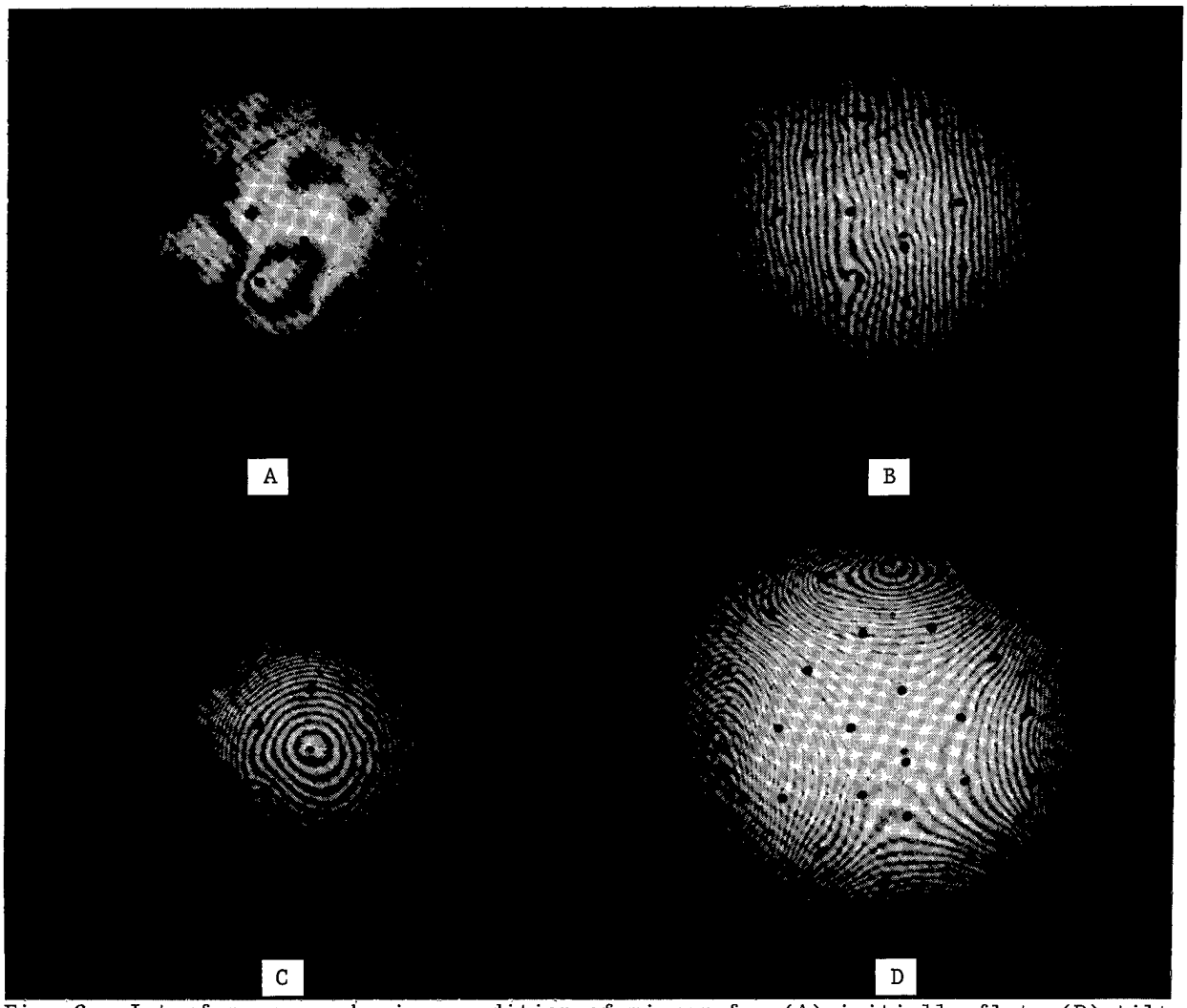


Fig. 6. Interferograms showing condition of mirror for (A) initially flat, (B) tilt, (C) focus, and (D) astigmatism.

The results of the quantitative analysis of the interferograms for this mirror are shown in Figure 7 for tip, tilt, focus, and astigmatism modes. The vertical axis represents actual mirror surface displacement, while the horizontal axis is the maximum applied voltage divided by a factor of 200 (i.e., 7.5 means 7.5 x 200 or 1500 volts). The solid lines are fitted to the data for the amplitudes of the desired modes. The data points indicated by an (X) are fitted to the RMS deviations of the actual surface from the desired mode shape. These results show good linearity between the mode amplitudes and input signals. The RMS deviation from the desired surface is usually between $\lambda/2$ and $\lambda/4$ at 0.6328 μm. Since this corresponds to $\lambda/32$ or better at the design wavelength, this mirror should exhibit very good performance when operated with any of these modes at the design wavelength.

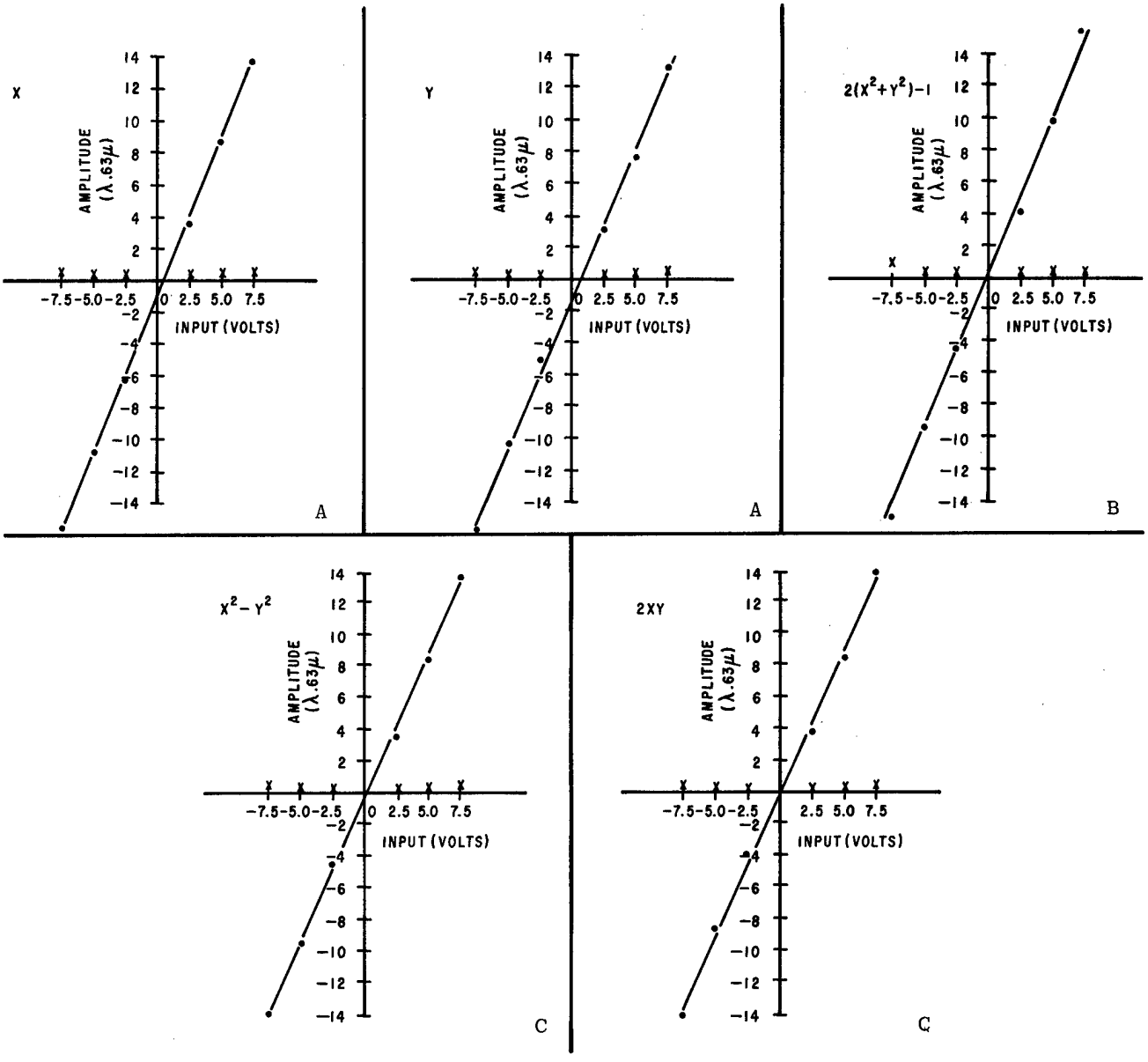


Fig. 7. Mirror performance for (A) tip and tilt, (B) focus and (C) astigmatism.

A computer program was developed that uses the measured influence function to calculate the expected mirror surface and the RMS deviation of that surface from the Zernike mode shape. Good agreement was obtained with the measured RMS deviation by summing the calculated RMS deviation of the expected surface response with the measured residual inherent RMS error of the undriven surface as uncalculated errors.

References

1. Kenemuth, J.R., SPIE Optical Components: Manufacture and Evaluation, 171, 32 (1979).

Computer-controlled optical surfacing

Allen H. Greenleaf
Itek Corporation, 10 Maguire Road, Lexington, Massachusetts 02173

Abstract

During the last 3 years under the High Altitude Large Optics (HALO) Technology Program , we have been developing computer-controlled, optical surfacing technology for application to the manufacture of large lightweight, odd-shaped, infrared-quality aspheric mirrors that will be needed in the next generation of space telescopes. The program has consisted of developmental work on our small optical surfacing machine, to support on-going production of a large oblong demonstration mirror with our large machine. The problems addressed have been the measurement of the aspheric surface, the material removal model, the data reduction machine control algorithm, and lap designs for producing aspherics rapidly. To date, flats and spheres have been produced with rigid laps, control of absolute sag to 0.1 μm has been demonstrated, and a compliant lap concept has been developed. Controllability of this lap has been verified in polishing. The proof test, to start shortly, will be a full grind/polish fabrication sequence to produce an eccentric ellipse. The program plan calls for later fabrication of a large two-panel eccentric ellipse, using the present large oblong demonstration mirror as the inner panel. A new concept for a self-calibrating, absolute measuring machine is being developed for use in the production of the large ellipse. It should be accurate enough for all of the grinding and perhaps for the polishing and final figuring too.

Introduction

Future generations of orbiting telescopes used for astronomy, information collection and transmission, and power transmission will require very large segmented mirrors with panels the diameter of the Space Shuttle cargo bay. The telescope designs will necessarily make use of fast highly aspheric mirrors in order to keep the telescope dimensions and weight to a minimum, and very thin glass sections will limit the usable grinding and polishing pressures. At first, optical surface tolerances will meet the requirements for infrared systems, but eventually mirrors with quality suitable for visible-spectrum use will be required.

Traditional optical fabrication techniques will be impractical for surfacing panels for these mirrors; highly automated methods will be needed. This general problem of automated surfacing of very large mirrors is one we are currently addressing at Itek, largely under the funding of the HALO Technology Program.

History of computer-controlled optical surfacing at Itek

Itek's first computer-controlled machine was completed in 1968 and was used for the initial research work in computer-controlled polishing. This machine is like a large x-y plotter and can place the lap spindle anywhere within a 50 × 50-in. square. The lap spindle rides up and down in bushings in the machine and is supported by the surface of the mirror being worked. In the earlier work, the surfacing energy was supplied by the machine as it translated the work over the surface (see Figure 1). With a conventionally sized lap, the machine could produce spheres of acceptable quality (see Figure 2) and had the advantage over a conventional machine of dealing readily with nonrotationally symmetric errors. In making aspherics, the machine had limited success. The lap for making an aspheric had to be smaller, and the process became very slow because of the increased ratio of mirror area to lap area. In addition, problems in dealing with the mirror edges were not completely resolved.

The opportunities afforded by the machine remained promising, nevertheless. It might permit the fabrication of surfaces that were not possible with conventional techniques—severe aspherics and odd-shaped mirrors. The process was inherently repeatable and self-operating, offering opportunities for a multiple-machine shop with a minimum of operators and multiple shifts for speeding up a production program.

A second, larger, computer-controlled machine, was built in 1970. One of the tests made with this machine was to figure an oblong mirror with a spherical figure.

PFR-80-028

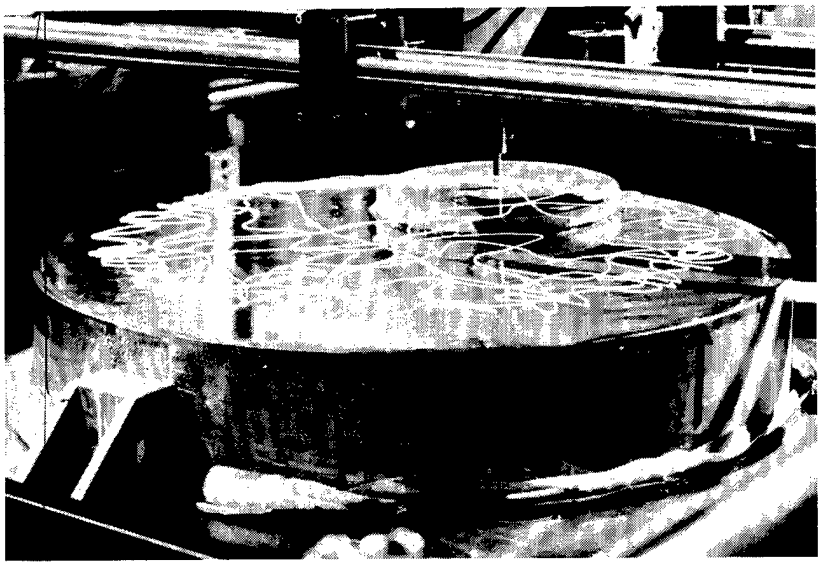


Figure 1. Tool movement on the 50-in. XY machine in its original configuration

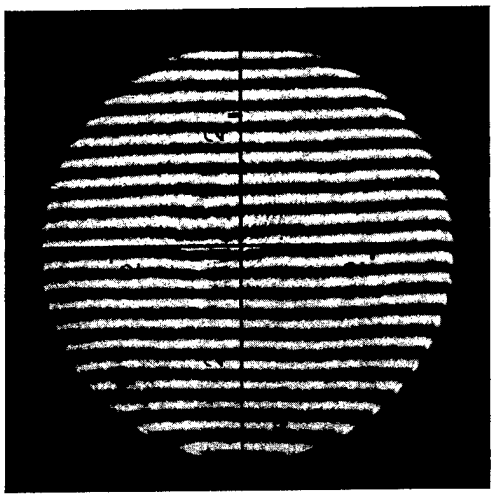


Figure 2. 24-in.-diameter spherical surface figured by the 50-in. XY machine

Present program

Objectives

The immediate objective of the present program is to support the on-going fabrication of a large oblong spherical demonstration mirror, but in a manner suitable for production of very large aspheric mirrors in the future. The surfacing system must be able to generate, fine grind, polish, and figure large aspheric surfaces, matching the radii of panels of segmented mirrors. The figure must be well-controlled close to the edges and the polishing must not produce undue waffling of lightweighted mirrors. In the interest of production of numerous, very large mirror panels, the surfacing time and measurement time must be minimized. As a practical matter, the surfacing process should not depend upon extremely high-density measurement data.

After the sphere is finished, production of a large, two-panel eccentric ellipse is planned, with the present oblong spherical mirror being refigured as the inner panel.

Table 1 gives the program objectives for upgrading the computer-controlled surfacing machines to a production system, and lists solutions that have been or are being implemented in order to meet the various objectives. In order to minimize working time, powered spindles have been added to the machines so that the surfacing energy comes from the spindle drive motors. The laps that have been used are annular laps with relatively large diameters in relation to the mirrors being worked (up to 1/2 the mirror diameter) so that the material removal rate is maintained at a high level. At the same time, the smoothing provided by a large lap tends to minimize the surface data density requirement in measurement and thereby simplify the measurement and data storage problem and speed up the measurement process.

Table 1. Aspheric Fabrication Problem/Solution Matrix

	Process Implementation						
	Powered Spindle	Maximize Lap Area	Controllable Lap Center of Gravity	Conforming Lap	Limit Polish Pressure	Extend Fine Grind	Develop Measuring Machine
Generate aspheric surfaces				X			X
Polish and figure aspheric surfaces				X			X
Match mirror panel radii							X
Control figure close to edge			X				
Prevent surface waffling					X		
Minimize surfacing time	X	X				X	X
Minimize measurement time		X					X
Minimize data-taking density		X					

Data reduction algorithm

The basic algorithm used in the computer program that converts surface data to machine commands is Preston's law, which assumes that the material removed at a given point on the mirror surface is proportional to lap pressure and the total linear motion of the lap across that point. The pattern of material removed from the mirror surface with the lap axis of rotation stationary is the lap work function. The computer uses an iterative algorithm to determine the amount of time the lap needs to spend in the vicinity of each site in an array covering the mirror in order that the surface error is minimized by the collective effect of

the lap work function applied at all the lap position sites. Two extra degrees of freedom of control are added by assuming that a gradient in the work function in any direction can be obtained at each site. This gradient would not be necessary for figuring an infinite surface, but with a finite surface it allows a good solution to be obtained at the mirror edges.

A means of controlling the position of the center of gravity of the lap has been added to each machine so that the desired gradient in the work function of the lap can be created. The center-of-gravity shift is accomplished by means of pneumatic cylinders that apply a tipping moment to the lap. The pneumatic pressure is under control of the computer through electric regulators. The lap mechanism is shown in Figure 3.

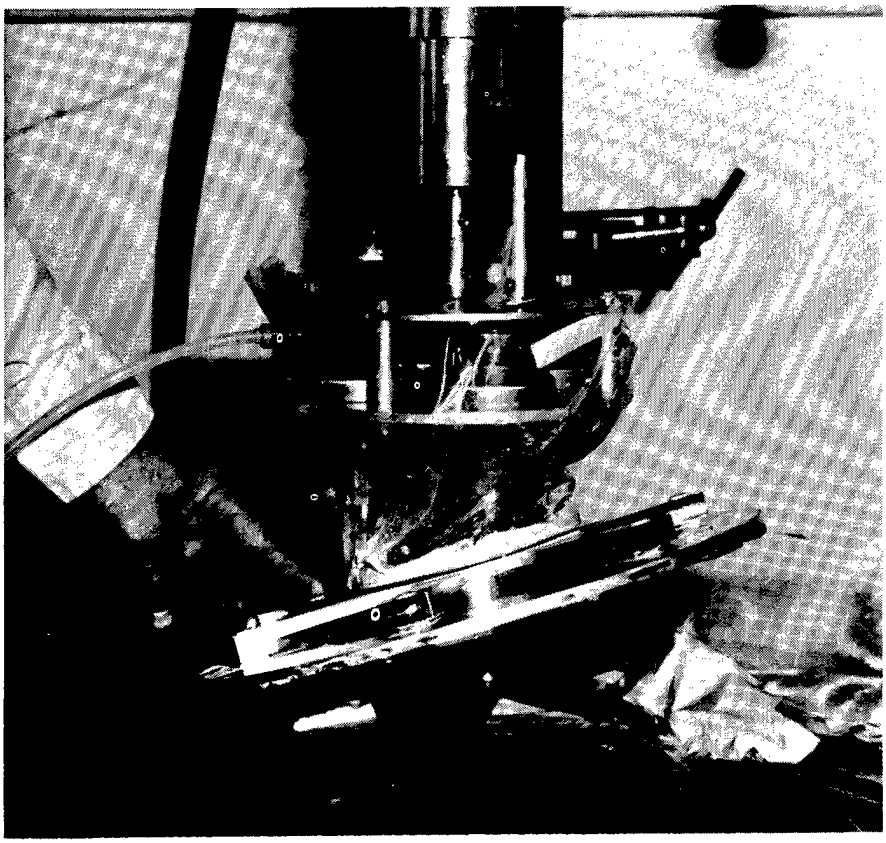


Figure 3. High-speed lap with controllable center of pressure

The algorithm in the computer model uses the assumption that the pressure between the lap and the mirror is independent of the local shape of the mirror surface or, in other words, that the lap is infinitely compliant. It probably would be impractical to take into account the lap stiffness, lap surface wear, and mirror compliance as would be necessary if the ability of the lap to iron out local bumps were to be modeled accurately. Consequently, for any given starting surface, the corrected surface predicted for the computer-selected lap time/position function and center-of-gravity function will contain may small irregularities associated with the lap site interval used in the solution. The actual lap control functions chosen, however, are probably not too different from those that would be selected if the modeling of the process were more exact, especially if it can be shown that the lap/mirror interface pressure variations caused by a misfit of the lap and mirror surfaces is but a fraction of the average interface pressure.

Laps for lightweight aspherics

In the case of the fabrication of spheres, misfit of the lap and mirror surfaces will tend to speed up the convergence of the mirror surface toward a sphere. In the case of the fabrication of an aspheric, the criterion for lap compliance that will permit convergence of the process to the desired aspheric without resorting to more complex modeling of the lap in the computer algorithm or an active lap that is forced to take the desired local shape is that the pressure variation at the lap/mirror interface caused by the misfit of the lap with the final, desired mirror surface be but a fraction of the average pressure. This is the approach being used in the present program for the surfacing of aspherics.

The approach retains some empirical results as an important ingredient in the selection of lap size and construction. Enough local stiffness must be retained in the lap so that it can perform its function of ironing-out the bumps, but the overall compliance must be such that the lap can fit the aspheric surface without excessive pressure variation. Fortunately, the effect of asphericity of the mirror for any real telescope, as far as the lap is concerned, is largely a change in the average curvature underneath the lap from place to place on the mirror or in a difference in the principal curvatures beneath the lap at any given location. If the lap can be designed to accommodate the greatest difference in average curvature within the mirror boundaries and the greatest differential curvature at any place on the mirror, then the lap should be adequate for aspherizing the mirror. An important factor in selecting the diameter of a lap is, therefore, the curvature variations of the mirror that must be accommodated by the lap compliance.

Another important factor in lap design and lap size selection is the average polishing pressure. The stiffness variations in the face of a mirror that are the result of a honeycomb or ribbed construction will cause local deformation underneath the pressure of a passing lap. This deformation causes an uneven material removal and results in a waffled surface of the mirror. The criterion for setting the maximum polishing pressure is the allowable waffling.

Assumptions about what a lightweight mirror a generation beyond the NASA Space Telescope primary mirror might look like indicate that polishing pressures should be limited to 0.1 to 0.2 psi. The pressure variation caused by the desired asphericity should be a fraction of that pressure, and that requirement coupled with a practical lap construction concept and the maximum curvature variations of the final surface constitutes an important limiting factor on the diameter of the lap. It is important to maintain the size of the lap in the interest of smoothing and total time for fabrication. The next advance in aspheric fabrication could include a combination of an actively deformed lap, more accurate modeling with the computer algorithm, and localized polishing support for the mirror.

As far as the lap is concerned, mirror asphericity is best expressed in terms of curvature variations, once the motions of the lap are placed under computer control and made infinitely variable. The aspherizing concept presented here indicates that once the lap is designed to the given criterion, the aspherizing of the mirror should proceed smoothly, and the tests made to date, which will be discussed shortly, have tended to verify this.

Surface measurement

The complement of the actual surfacing is the measurement of the surface to provide the data for controlling the surfacing machine. The development of a measuring machine to provide surface measurements as accurate as possible is an essential part of the program to develop a system for the surfacing of large aspherics. The measuring machine has the important function of providing all the surface data for the fabrication process, at least to the point at which good interferograms can be made. If the measuring machine can be made accurate enough, then even the interferograms, which involve mirror-specific special test optics and, in the case of very large mirrors, very large working distances, can be eliminated. Another important function of the measuring machine is to measure radius differences between two panels in a segmented mirror so that the surfacing machine can match the panel radii. All of the measurements must contain the absolute radius error, because the radius error and the higher-order figure error of a mirror panel must converge to tolerance together.

A good measuring machine will reduce the measurement part of the fabrication cycle. The more accurate the measuring machine, the more accurate will be the surface grinding, and so the polishing time can be reduced. Also, the more accurate the measuring machine, the greater will be the benefit from using harder, finer abrasives in the fine grinding, and as the grinding is extended to finer and finer grits, the less will be the material removal required in polishing for the elimination of the subsurface damage left by the grinding. The measuring machine therefore becomes an important factor in minimizing the surfacing time, even if its accuracy is not sufficient for final figuring. Measurement techniques that have been used to date and the measuring machine concept now under development will be described shortly.

Program plan and status

The present computer-controlled optical surfacing development program is outlined in Figure 4, and its current status is indicated. The program began with the modification of the two surfacing machines to include the powered spindles and the center-of-gravity control for the laps. The 50-in. machine in its present configuration is shown in Figure 5. The machines originally were driven from commands on tape through hard-wired controllers. To provide more flexibility in control agorithms, the hard-wired control systems were replaced with a computer that controls the machine stepper motors, pressure regulators, and spindle

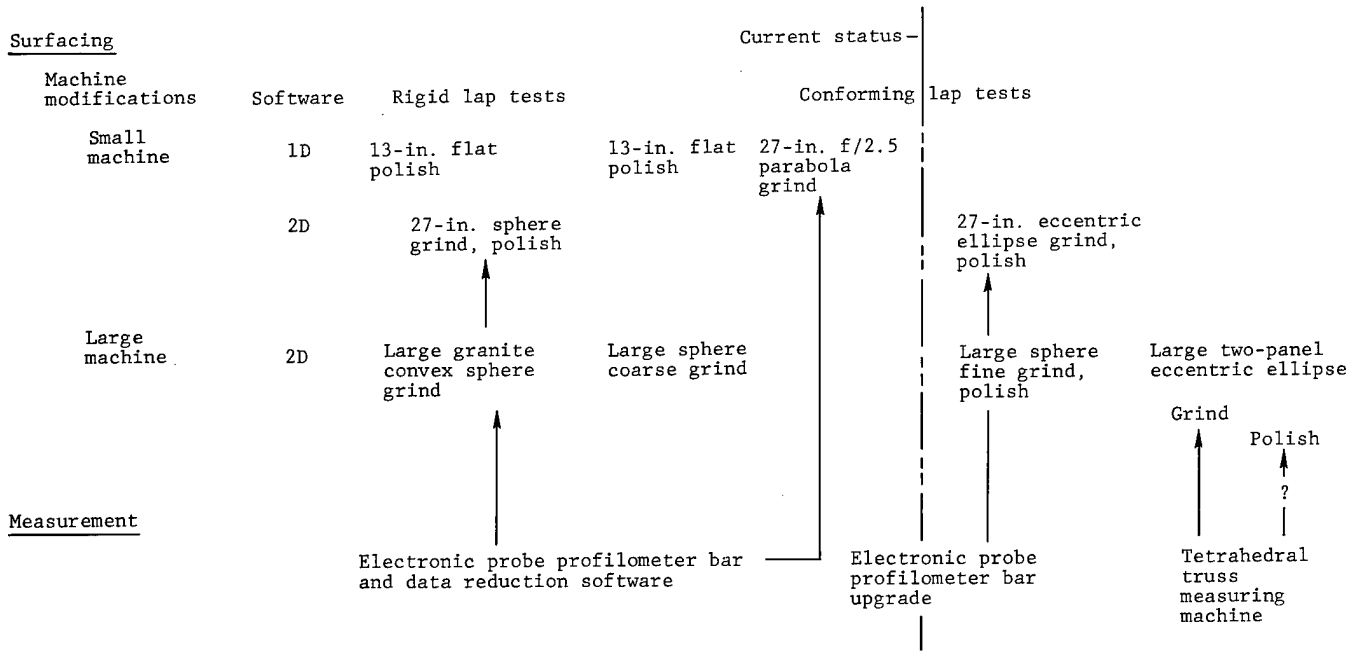


Figure 4. Computer-controlled optical surfacing development program

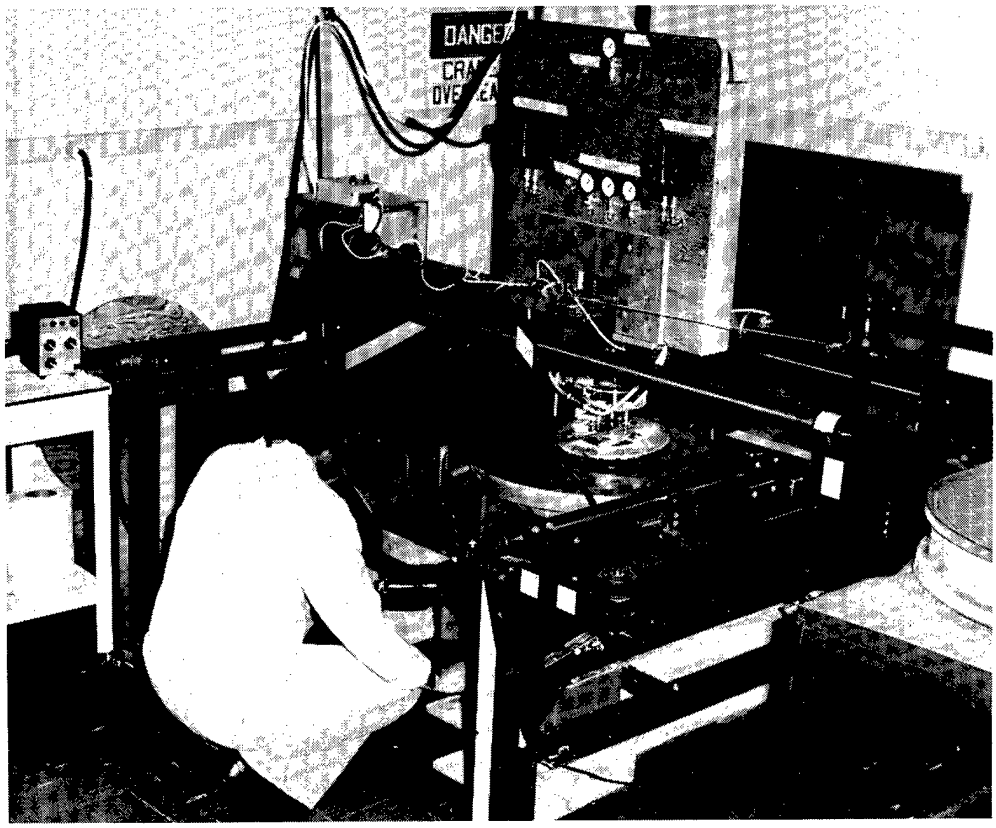


Figure 5. 50-in. computer-controlled optical surfacing machine

directly. On-going development of the control computer software will soon give us the capability of operating two or more machines simultaneously with the same computer.

One-dimensional control option. Two forms of the data reduction software were written. One uses only the rotationally symmetric errors in the measured surface and generates machine commands to move the lap and to move the lap center of gravity in radial-only direction with respect to the mirror. A turntable was added to the smaller of the two surfacing machines for use with this one-dimensional data reduction program. Preliminary concept tests are made on the small machine operating in the one-dimensional mode. The advantages of the simpler (radial only) surface measurement, data reduction, and machine operation have resulted in significant savings in experiment time. After the testing and debugging of a lap concept or algorithm modification has been completed in the one-dimensional mode, the experiment can be repeated in the two-dimensional mode, in which the lap and the lap center of gravity are controlled in two directions and the rotating table not used. The larger surfacing machine does not have a rotating table at present and has, in fact, generally been used for oblong pieces for which a turntable would not be useful.

Rigid lap tests. The significant tests made to date and the tests planned are shown in Figure 5. The modified machines and data reduction algorithms were first tested with rigid laps. A polishing lap used to figure a 13-in. flat in the one-dimensional mode is shown in Figure 6. Figure 7 shows the surface of this flat after the initial polish from the ground condition and after the final polish, five iterations later. For this surface, the circumferentially averaged rms residual error was 0.035 μm rms with respect to the test flat.

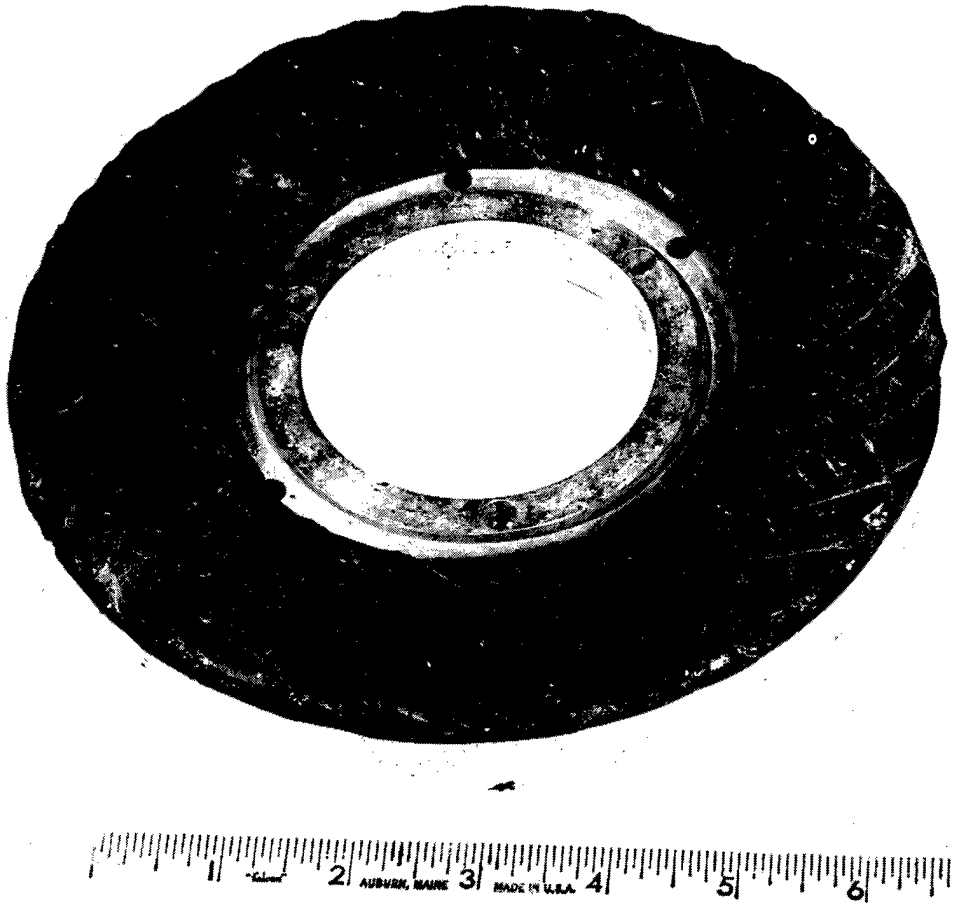
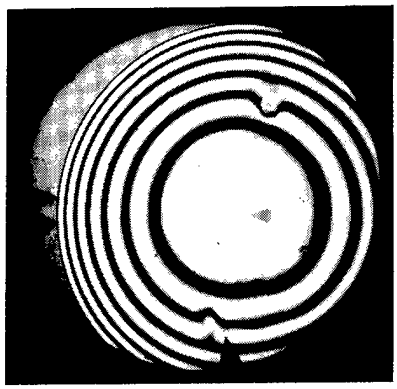
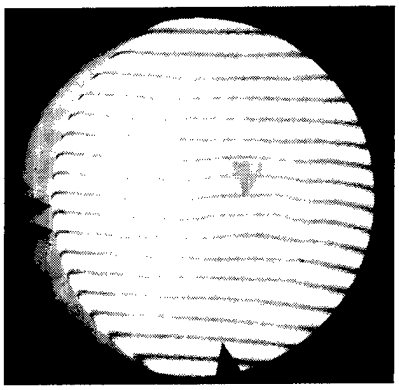


Figure 6. 7 3/4-in. rigid pitch lap used to polish 13-in. flat



Before: 13-in. flat with 0.66 μm rms of focus – tested relative to reference flat



After: 13-in. flat polished on small computer-controlled optical surfacing machine to 0.035 μm rms – tested relative to reference flat.
Machine time: 8.4 hours

Figure 7. 13-in. flat polished with rigid lap

After polishing the 13-in. flat with the one-dimensional control, a complete grind/polish sequence was used to figure a 26-in. sphere with two-dimensional control. The measurements for the grinding of the sphere were taken with a profilometer bar with ten electronic probes. The outputs of these probes were fed directly into a computer, and data reduction software written for the profilometer-generated surface errors from the network of bar positions used and probe readings recorded. Figure 8 shows the profilometer bar being used to test the mirror in situ on the surfacing machine. Figure 9 shows the bar placement pattern and the probe locations used for the measurement. Figure 10 shows the computer-reconstructed error surface at the start and at the end of the grinding process. The reference surface used for zeroing the probes was another sphere, and the surface sag match at the end of the grind was within the 0.1 μm resolution capability of the bar/probe system. After grinding was completed, the mirror was polished with a 12-in.-diameter rigid pitch lap similar to the one shown in Figure 6, and the figuring was stopped at the result shown by the interferogram in Figure 11. At this point, the rms surface error was 0.045 μm, and the surface sag, as measured with the profilometer bar, still matched that of the reference sphere to within 0.1 μm.

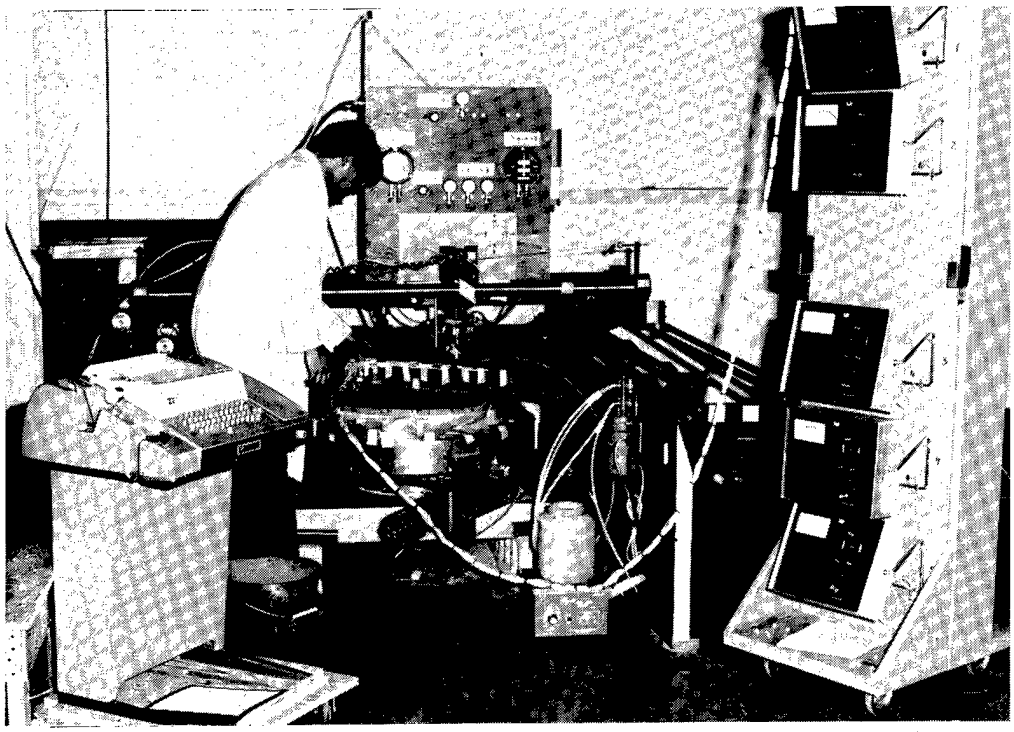


Figure 8. Profilometer bar being used to measure 27-in. sphere in place on 50-in. surfacing machine

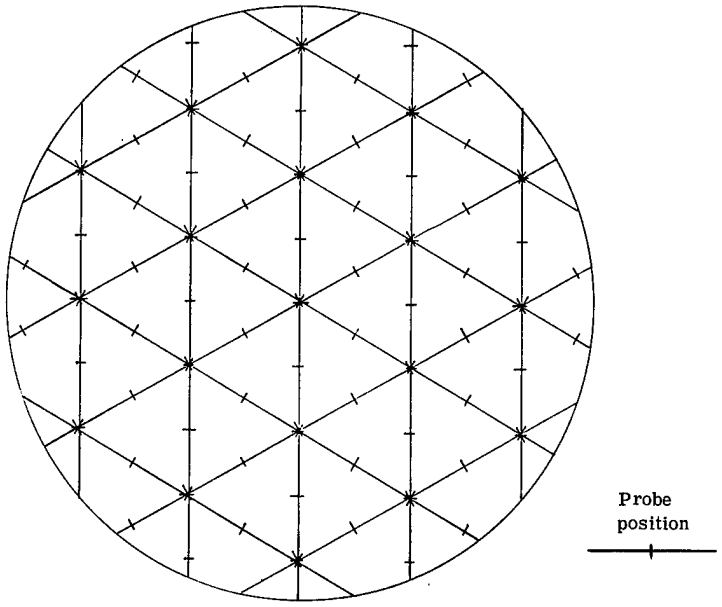
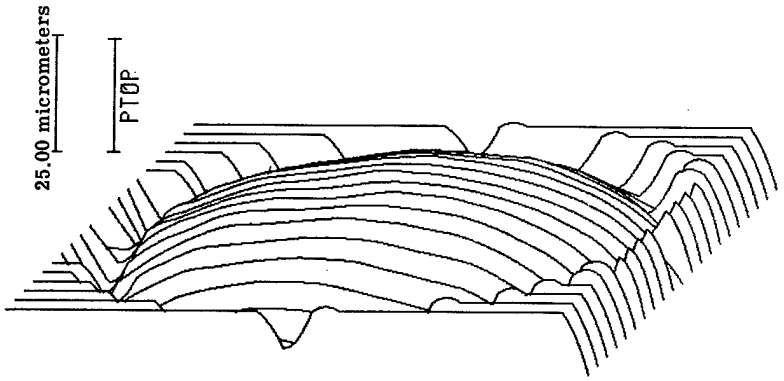
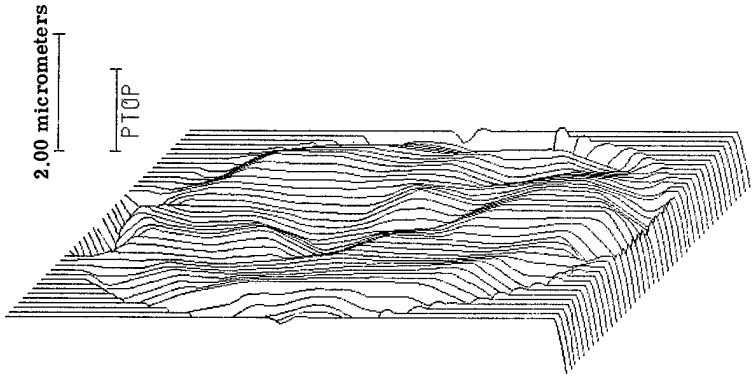


Figure 9. Profilometer pattern for 27-in. sphere



(a) After first grind with 3 μm abrasive — 1.4 μm rms, 14 μm sag error



(b) After seventh grind with 12 μm abrasive — 0.18 μm rms, 0.1 μm sag error

Figure 10. Profilometer measurements of 27-in. sphere

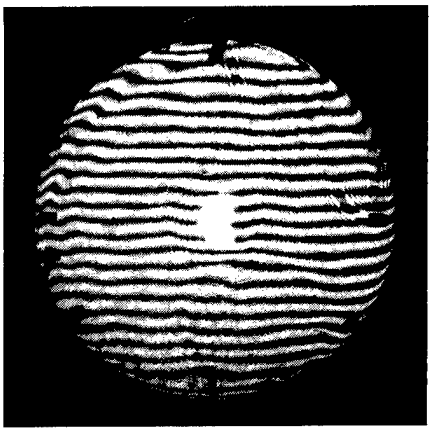


Figure 11. 27-in. sphere polished with rigid lap, 0.20 μm p-p, 0.045 μm rms, 0.1 μm sag error, 44 hours effective machine time

Conforming lap tests. After the tests with rigid laps, tests were begun using a conforming lap designed according to the principles previously discussed. It was necessary first to make an assessment of what curvature variations would represent real-life aspherics of the future. Curvature has dimensions of reciprocal length. A survey of potential very large telescopes and the aspherics that might be used in them indicated that curvature variations on the order of 0.002 to 0.01 per meter would have to be accommodated by the grinding and polishing laps.

A 6-in. diameter polishing lap, shown in Figure 12, was designed to accommodate a 0.01 per meter curvature variation. The first question to be resolved regarding this lap was whether the flexibility introduced would result in the loss of the smoothing capabilities of the lap and loss of figure control at the edge of a mirror. This lap was used, therefore, to polish the 13-in. flat again from a ground state using the one-dimensional control algorithm. At the stopping point for the test (see Figure 12), the circumferentially averaged error was 0.025 μm rms. Next, an 11-in.-diameter lap was designed and built (see Figure 14) to grind and f/2.5 parabola into a 27-in.-diameter blank. The curvature variations for this parabola are about 0.004 per meter. The measurements during this test were taken with the profilometer bar in four radial positions to produce 40 radial data points for the input to the one-dimensional data reduction algorithm. Figure 15 shows the theoretical departure of this parabola from the best-fit sphere, the actual initial profilometer measurement, and the profilometer measurement taken after the work was stopped. The final result was a 0.36 μm rms surface, generally smooth, and a good starting point for the polishing phase of the fabrication sequence.

Preparations are now being made to grind, polish, and figure an eccentric ellipse on a 27-in.-diameter blank, using the two-dimensional data reduction algorithm. The curvature variation for this ellipse is about 0.004 per meter. A locating fixture is being made for the profilometer bar so that its position on the surface will be accurately known.

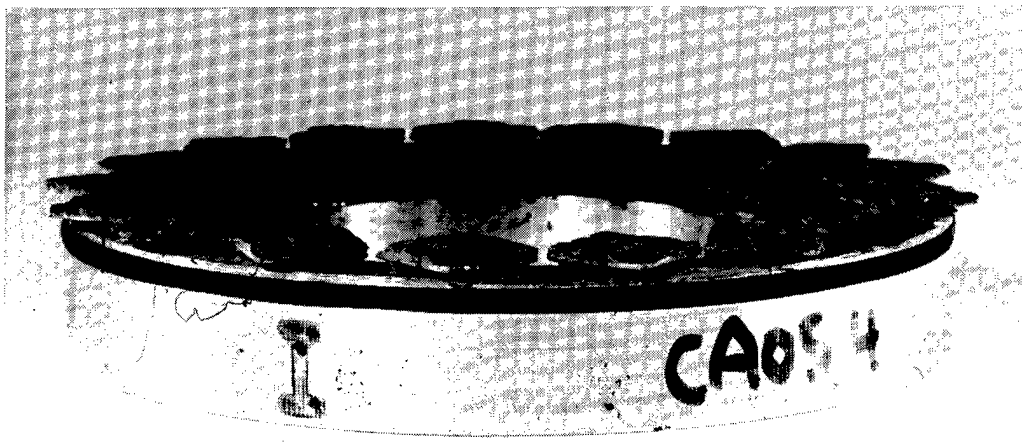


Figure 12. 6-in. compliant polishing lap used to polish 13-in. flat

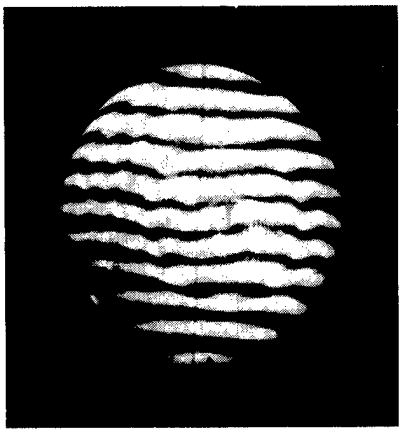


Figure 13. 13-in. flat polished with compliant lap. Total polishing time 5.6 hours, 0.025 μm rms (circumferential average)

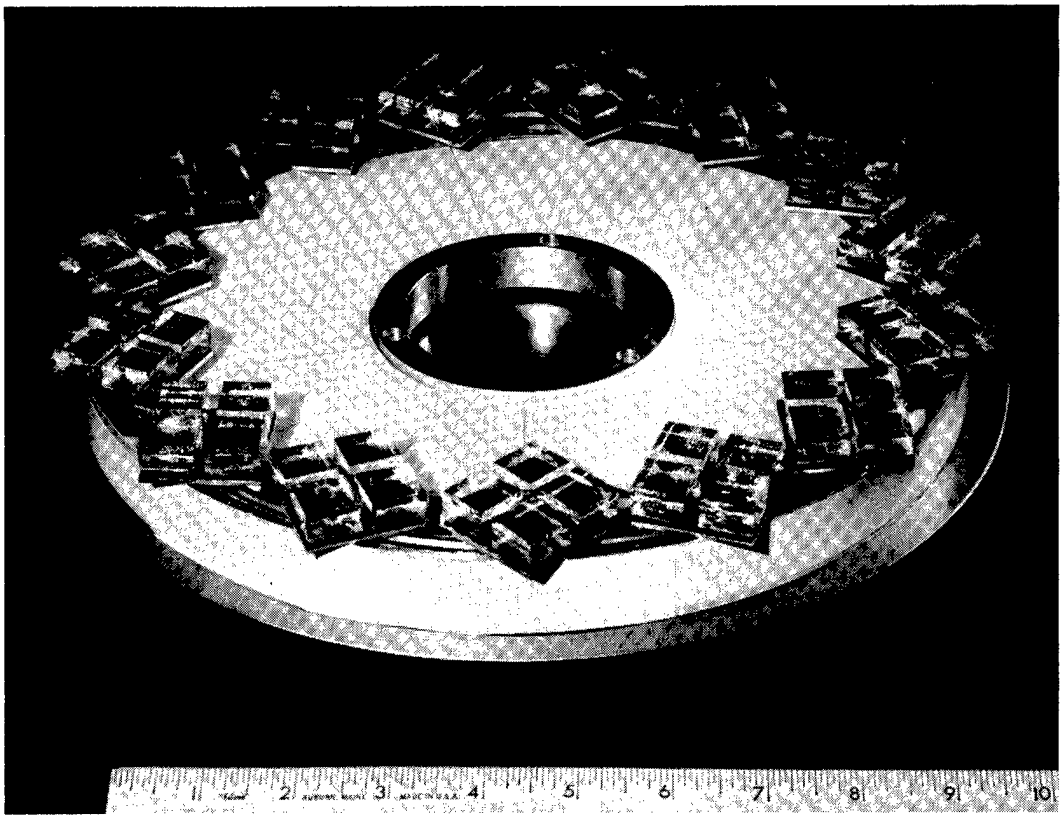
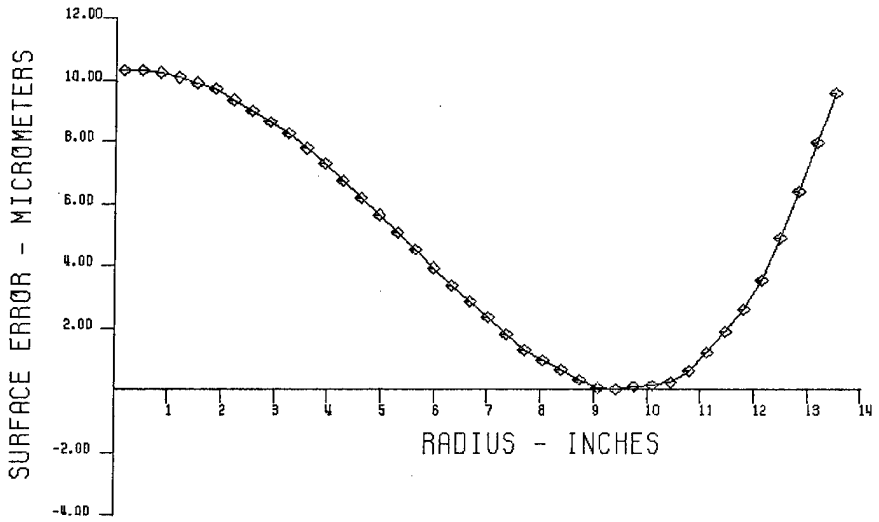
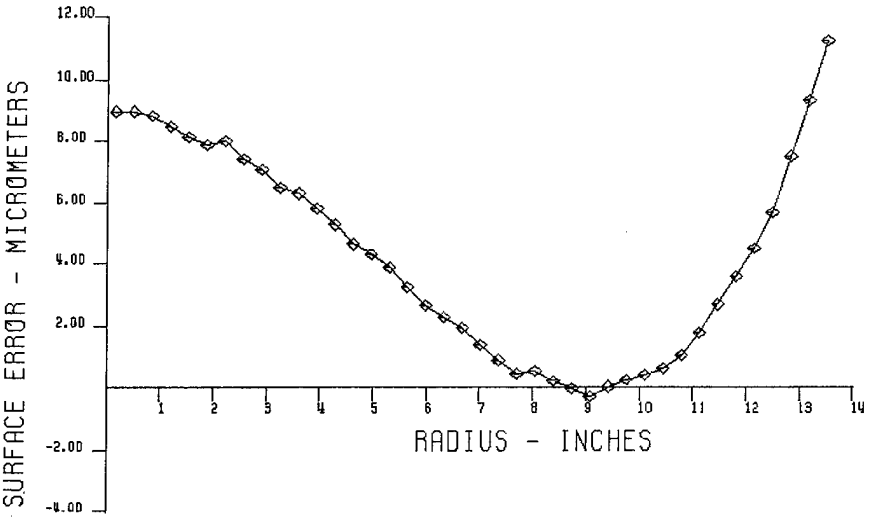


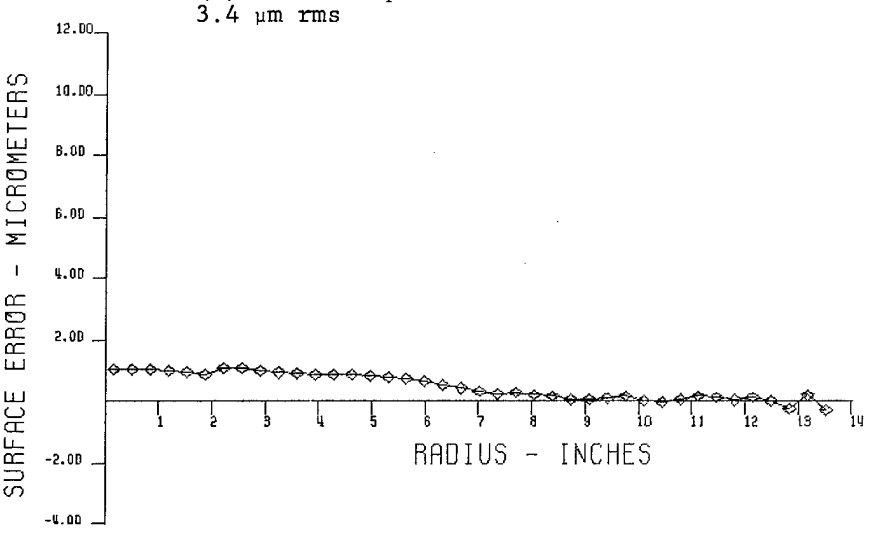
Figure 14. 11-in. compliant lap used to grind 27-in. parabola



(a) Removal required from best-fit sphere



(b) Initial profilometer measurement
3.4 μm rms



(c) Initial profilometer measurement 1.24 hours machine time
5 μm grind 0.36 μm rms error

Figure 15. Parabolization of 27-in. f/2.5 mirror

Work done with large surfacing machine. The large surfacing machine, after modification to include the powered spindle and the variable center-of-gravity capability for the lap, was used to grind a large slab of granite, with a convex spherical surface as received from the quarry, to an error of 0.0005 in. rms. An average of 0.1 in. of granite was removed from the surface. The granite slab is tooling for an unrelated test.

The program plan includes production of a large two-panel, eccentric ellipse. The large surfacing machine was used to generate spherical surfaces on both sides of the first of the oblong blanks for this mirror. The next stage will be to grind and polish a spherical figure on the front side of this blank, which will become the inner and least-aspheric of the two panels. Then the asphérization will be started, first with the inner panel, then with the outer panel.

Measuring machine

The profilometer bar has performed very well in the uses made of it, but its usefulness for measuring large, severe aspherics is limited by the linearity limits of the probes, and its ultimate accuracy probably cannot be improved much beyond that which was achieved in the measurement of the 17-in. sphere. Itek is currently developing another measuring machine concept that we plan to use for measurement of the second panel of the eccentric ellipse.

One approach to measuring a convex or concave surface is to triangulate to a retroreflective target on the surface from a set of observation points above the surface. Triangulation from three points would be sufficient if the distances between the three points and the distances from the three points to a single point on the surface were accurately known. Adding a fourth observation point to form a tetrahedral arrangement above the surface, as shown in Figure 16, gives a redundant measurement. As the retroreflector is moved above the surface and the changes in the distances from the observation points are measured, enough redundant data can be taken so that the dimensions of the tetrahedron and its position relative to the measured surface can be calculated. By counting the measurements and the unknowns it appears that after 10 measurements, all of the necessary dimensions can be determined. This method appears to offer, therefore, a relatively low-cost approach to using a system of differential distance-measuring interferometers (such as the Hewlett-Packard) for absolute surface measurement.

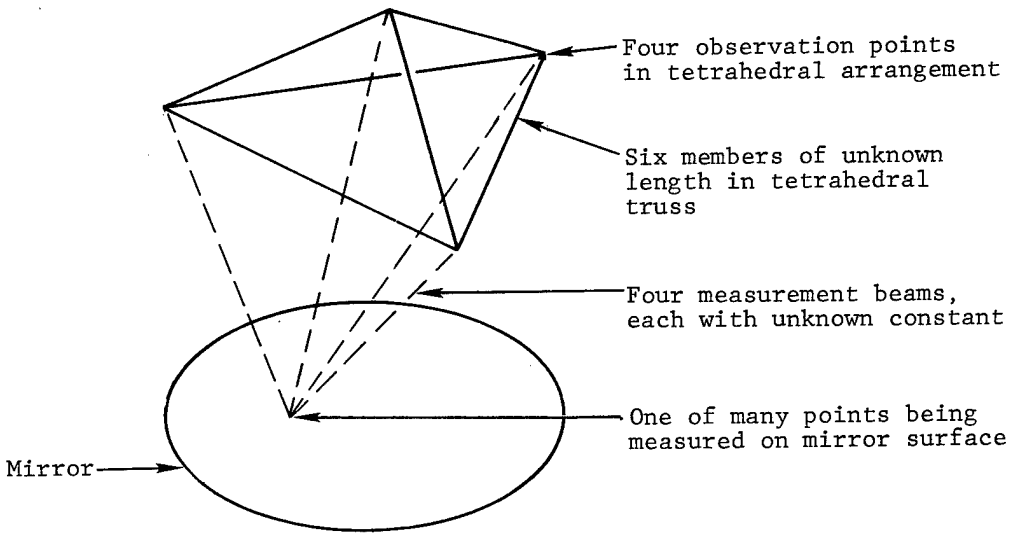


Figure 16. Self-calibrating absolute surface measuring system

During the course of the measurement process in which the retroreflector is moved about the surface, the dimensions of the tetrahedron must either be maintained constant or monitored for changes. Similarly, the position of the measured surface relative to the tetrahedron must be maintained constant or monitored. These dimensions will be fixed by using graphite-epoxy struts and metering rods. A single laser, serving steerable interferometers at the four observation points, will be used to make the measurements to the retroreflector on the measured surface.

Although, theoretically, the system unknowns can be determined after 10 measurements, errors in the determination of the system parameters resulting from noisy measurements will continue to diminish as more and more data points are added to the surface. A normal measurement used for controlling the surfacing machine would probably consist of an array of points with 50 to 100 points across the diameter, or between 2,000 and 8,000 points on a round mirror. A computer simulation of this machine has indicated that after about 200 points have been measured, the surface error has about the same rms value as the rms error in measurement of a single interferometer beam. This machine, when completed, should form an extremely important part of the overall surfacing system, perhaps even eliminating the need for center-of-curvature testing of the final figure and making the fabrication of large convex aspherics as straightforward as the fabrication of concave pieces.

Conclusion

Under the HALO Technology Program, a computer-controlled optical surfacing system is being developed for rapid surfacing of very large, lightweight, highly aspheric mirrors of infrared quality. Although we have not yet been through the full fabrication sequence with a severe aspheric, the experiments completed indicate that our process operates in the intended manner. Fabrication of the eccentric ellipses will be the proof tests.

Semiautomated measurement with the profilometer bar in the grinding phase has proved to be sufficient in preparing the surface for interferometry and polishing. With the profilometer bar, sags of separate pieces have been matched to 0.1 μm.

The measuring machine under development promises to improve the surface measurement accuracy in the grinding phase even further, and may eventually be able to perform some or all of the figuring measurements now made with center-of-curvature interferometry, eliminating test problems associated with aspherics and convex surfaces.

Acknowledgements

The development work leading to this paper has been done largely under the High Altitude Large Optics (HALO) Optical Technology Program, contract No. F3060-76-C-0336 of the Defense Advanced Research Projects Agency. Preliminary analysis of the tetrahedral truss measuring machine concept has been performed under the Advanced Optics Fabrication Technology program, Contract No. F30602-79-C-0251 of the same agency.

ACTIVE OPTICAL DEVICES AND APPLICATIONS

Volume 228

SESSION 2

TECHNOLOGY FOR LARGE ADAPTIVE OPTICS II

Session Chairman
Thomas G. Pitts
Rome Air Development Center

Integrated sensing and control system for a large, deployable, wide-field optical system

John T. Watson, Dennis C. Ehn
Itek Corporation, 10 Maguire Road, Lexington, Massachusetts 02173

Abstract

Large lightweight deployable optical systems are highly sensitive to structural changes from material aging and from solar heating. Optical tolerances are a fraction of a wavelength of the radiation the system is designed for and the tolerances remain constant regardless of system size. These large three-mirror infrared systems are to be manufactured at normal temperatures and for use at cryogenic temperatures. Even the most uniform optical material known, fused silica, becomes one of the quality limitations at these temperatures. The quality maintenance problem is compounded when many mirror panels must be assembled to provide a single large mirror area. The support structure must keep individual panels accurately located on one mathematical surface to a small fraction of a wavelength of light. The structural materials that are available to mount the mirrors are an order of magnitude poorer in stability than the fused silica mirror panels, thus large structural warpages with time and with the position of the sun must be accommodated by system actuators between the mirror panels and the support structure. A sensing system is described that measures mirror panel-to-panel mismatch and that determines system wavefront as a function of sensor location in the image field. These data are accepted by a central computer control system that deconvolves the control signals. Examples are given of computer simulation of the sensing and control process, showing the number of iterations required to bring a system into optical adjustment.

Introduction

Proposed space telescope systems for the future include large astronomical telescopes for reaching farther into space, large mirror systems for transmitting collected solar energy to the earth, and systems for beaming power to satellites from a space station. High-quality systems with large mirror diameters are being envisioned. Among the technical problems encountered is that mirrors of this size are much too large to be made in one piece. Such mirrors will, therefore, have to be composed of many smaller mirror panels accurately held together to optical tolerances.

A study for a 10-meter, earth-based telescope being conducted at the University of California at Berkeley[1] conceives the mirror to be made up of 55 individual mirror panels. Panel-to-panel position errors are to be sensed by capacitance gauges.

Other organizations have been studying spaceborne telescopes. One candidate design, shown in Figure 1, is a three-mirror system with no central aperture obstruction and a wide field of view. It can be noted that since this system has appreciable field, the image-forming beam uses only part of the primary and tertiary mirrors for a given image point. The mirrors, as in the Berkeley telescope, are too large to be made in one piece. The proposed segmentation of the primary, secondary, and tertiary mirrors is shown in Figure 2. In the figure, the circles on the primary and tertiary mirrors show the portion of these mirrors that is used for imagery at the edge of the field. The complete surface of the secondary mirror is used since it is the stop of the optical system.

When one forms an image through such a system, image-forming wavefronts show the superimposed divisions of all the mirror areas used to form the image. Typical wavefront divisions are shown in Figure 2 for an image located at the edge of the field.

To obtain full aperture resolution in a segmented mirror system, adjacent panels must be controlled in position so that there is no "step" discontinuity in the surface figure of the complete mirror. This phase adjustment must be maintained to a small fraction of a wavelength of light.

The structure for such a system must be lightweight because of the space application. Furthermore, it should be of a low-expansion material so that changes in radiant heating, either from the sun or from the earth, will not cause rapid thermal warping of the large members. The structure must hold mirror spacings and alignments stable, generally to a few micrometers for spacings and to a few tens-of-micrometers for decenters and tilts. However, even with the best structures we can conceive, the structure may be able to hold acceptable

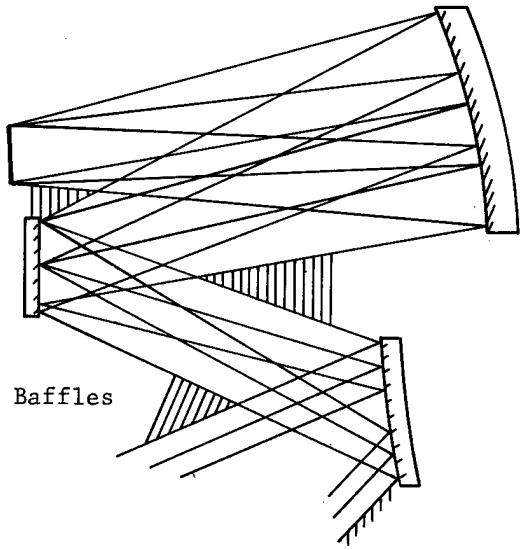


Figure 1. Three-mirror f/3 configuration

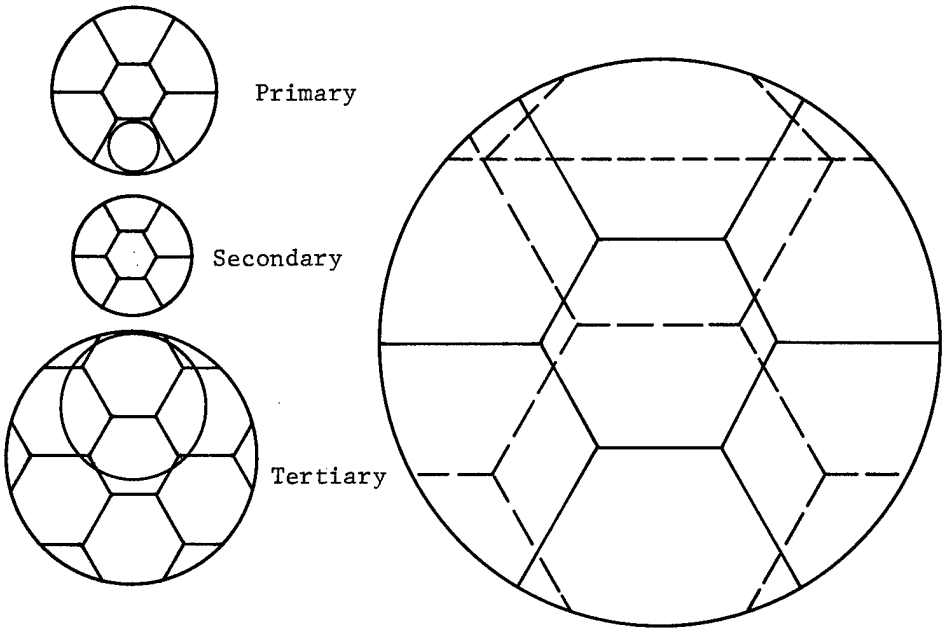


Figure 2. Mirror and pupil segmentation

alignment for only a short period. Thus, the system needs an error-sensing and system-readjustment capability.

Because of weight limitations, normal astronomical mirror configurations are not usable. To save weight, ultralightweight mirror panels could potentially be made from a thin facesheet of fused silica, supported by a stiff graphite-epoxy substrate. Because graphite-epoxy is not nearly as stable as fused silica, it is expected to warp, with temperature and age, well beyond optical tolerances. For this reason, the facesheet cannot be bonded directly to the substrate. It must, instead, be supported from the substrate through connectors whose length can be adjusted. In this way, minor warping of the graphite-epoxy substrate can be accommodated once the deformation of the mirror surface has been determined.

With this introduction, a look at how such a system might be set up and aligned is in order.

Deployment

Structures for systems of the type described are too large to be built on the ground and sent into space in their operational configuration. Potential assembly approaches range from having the structure assembled from parts in space to having it automatically deployable. It appears most efficient to use an unfoldable structure and to mount mirror panels and sensors on this structure after deployment. Even with the most carefully made structure and fittings, the structure as deployed in space is expected to have dimensional errors on the order of a millimeter. Thus, when the system is finally assembled in space, dimensional errors of the system are expected to be large compared to the tolerances required for good optical performance.

Panel-to-panel error removal

The first step after deployment of the system will be to compose the individual mirror panels to make primary, secondary, and tertiary mirrors that can give coarse imagery at the focal plane. The individual panels would be adjusted to make complete mirrors as follows:

1. Individual electrical contacts would be provided at points along the back edges of the panels so that the panels can be adjusted for proper orientation to one another. The electrical contacts will be in the form of "go/no go" gauges that will indicate when the panel-to-panel position, lateral and axial, is correct within about 25 micrometers. For lateral positioning, 25-micrometer accuracy is sufficient even for the assembly of strong aspheric surfaces. However, the axial panel-to-panel position must be controlled to less than 0.1 micrometer to prevent significant phase "steps" in the image-forming wavefront.

2. For "phasing" the panels to a fraction of a wavelength, there are several methods that could be used. For unattended space use, the phase-sensing devices would probably be grating autocollimators (Figure 3) that employ pairs of small mirror facets mounted on the back edges of adjacent mirror panels. Such devices are capable of controlling phasing to 0.01 wavelength and in addition, can control panel-to-panel tilt to about 0.1 second of arc. Moreover, the grating autocollimator is not at all sensitive to small errors of its alignment with respect to the mirrors.

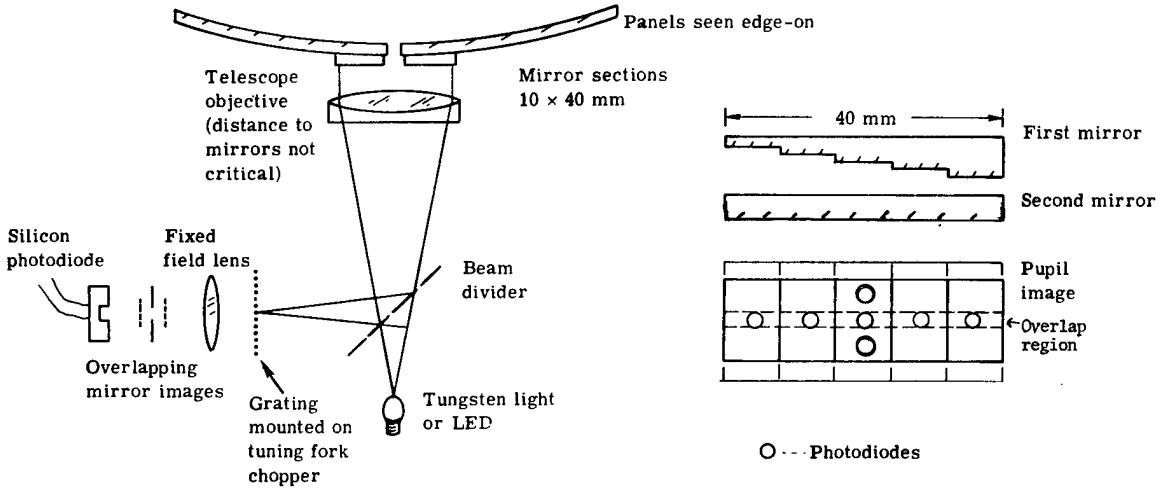


Figure 3. Grating autocollimator

Thus, the mirror panels will be adjusted to make reasonable primary, secondary, and tertiary mirrors, using internal error sensors and position actuators for removing the errors. For this purpose a sensor group, consisting of contact sensors and a grating autocollimator, would be located at the center of each division between panels.

System alignment and figure correction

The next stage in system refinement must be accomplished through the use of wavefront sensors. The wavefront sensor proposed is one that analyzes images from faint stars to provide wavefront shape. The sensor is a shearing interferometer, with light intensification so that individual photon distributions can be sensed above detector noise. The sensor determines wavefront slope departures from the correct spherical wavefront as a function of location in the pupil. Complete descriptions of this type of sensor and the signal processing can be found in the literature.[2,3]

If the entrance gate of the sensor is about 20 millimeters in dimension, the largest aberrated point image that would fit within the gate and provide light from all portions of the imaging wavefront is 20 millimeters. Since a star image would not generally be centered in the gate, only partial wavefronts will be measured for star images 20 millimeters in dimension. However, even partial wavefronts could be used to indicate what is wrong with the alignment and allow improvement to be made.

One system control concept calls for a complement of 12 wavefront sensors equally spaced around the edge of the focal plane, and one sensor in the center. The wavefront sensors allow determination of three types of error in the optical system. These errors are: (1) panel-to-panel residual tilt error, shown by wavefront slope discontinuities across the boundary between panels; (2) primary, secondary, and tertiary mirror decenter, tilt, and despace errors, shown by the coma, astigmatism, spherical aberration, and focus errors; and (3) surface figure errors caused by warpage. These errors are all summed together in the wavefront information obtained from the wavefront sensors.

Separating the errors and defining the best way to correct them is done by computer analysis of the data. Discontinuities in the slope of the wavefront surface come from only one source — panel-to-panel tilt errors. Discontinuities can be recognized in the data and correction applied.

System-alignment error and surface-figure error can each cause coma, astigmatism, spherical aberration, and focus errors. In fact, if the system were out of alignment, a reshapement of all the panels could theoretically provide complete correction. Analysis can determine the best way to make a good system, either to align the system or to warp the panels. The procedure would be to do as much of the correction job as possible by rigid body motions of primary, secondary, and tertiary mirrors. Residual errors that cannot be corrected by the rigid body motions would then be dealt with by warping individual panels.

There is still a question of which panels to warp since each wavefront contains the summation of the errors produced by three separate mirrors. However, data from the 13 wavefront sensors previously mentioned are sufficient to answer this question. The influence of each surface actuator on each of the 13 wavefronts is stored in computer memory. Mathematical analysis will indicate which actuators to use to optimize all 13 wavefronts simultaneously. The following section discusses computer simulation of the system control process.

Computer data analysis and system control

Data analysis method

The foregoing discussion has outlined the configuration of a large, multipanel, multisurface optical system. The wavefront and edge-displacement sensors are to provide the data for control of the system. The natural question is whether these data are sufficient to provide the desired degree of alignment and figure control. We shall describe a method that allows us to answer this question and aids in the diagnosis of potential weaknesses.

First we note that after the deployment and initial setup, the system is quite close to the nominal design. The relationship between actuator movement and observed error is linear. Thus it is appropriate to base the analysis on the response of the nominal system. This is standard practice in the analysis of manufacturing tolerances in optics.[4] The linearity has been verified over large ranges of system perturbations.

The sensor measurements are related to the actuator settings by the system observation matrix

$$Hx = z \tag{1}$$

where z represents a vector containing the full set of measurements and x is a vector of the actuator settings. The system observation matrix, H, is derived from rays traced backward through the system from the wavefront sensors and from the surface displacements produced by the actuator movements. If the observation matrix is sufficiently well conditioned, the required actuator setting changes may be obtained from the wavefront and edge sensor measurement by solving Eq. 1. (Often it is found that detailed analysis of the system equations simply confirms intuition.)

Singular value decomposition (SVD) is a convenient tool for the analysis of linear systems, particularly when numerically stable algorithms are available to perform the decomposition.[5,6,7] Any matrix, including the observation matrix, may be decomposed into

$$H = USV^T \tag{2}$$

where S is a diagonal matrix containing the singular values and U and V are orthogonal

$$U^tU = 1$$

$$V^tV = 1 \tag{3}$$

Substituting Eqs. 2 and 3 into Eq. 1, we may write the system equation:

$$S\,V^t x = U^t z \tag{4}$$

The singular value decomposition form of the system observation matrix equation facilitates analysis of the properties of the solution. Let the vector g be defined:

$$V g = x \tag{5}$$

Then the system equation becomes

$$S g = U^t z \tag{6}$$

The components of g represent the contributions of the various modes of actuator settings to the solution. Each mode is the linear combination of actuator settings specified by a column of the V matrix. The solution for each component of g is simple because S is diagonal.

$$g_i = (1/s_i)(U^t z)_i \tag{7}$$

or in the matrix form

$$g = s^{-1}U^t z$$

Noise analysis

The solution given in Eq. 7 is the "naive" least squares solution for the actuator settings given the measurements z. The actual sensors will have a certain amount of noise, n, in the measurements. The actual measurement, z', is given by

$$z' = z + n \tag{8}$$

For our discussion, we shall assume that the measurement errors are uncorrelated with equal variance, N^2. The error in g, dg is obtained using Eq. 7.

$$dg = s^{-1}U^t n \tag{9}$$

The error variance for an actuator mode is then given by

$$\langle (dg_i)^2 \rangle = N^2/S_i^2$$

The error for the actuator forces is given by

$$\langle |dx|^2 \rangle = \langle |dg|^2 \rangle = N^2 \text{ sum } (1/Si^2) \tag{10}$$

The size of the singular values is a measure of the degree to which linear combinations of actuator settings, the modes, can be observed and controlled. Part of the analysis of a particular system configuration then is just a matter of obtaining the system of equations

giving the sensor response as a function of the actuator settings, and performing a singular-value decomposition verifying that all the singular values are of similar size. In the evaluation, those linear combinations of actuator setting corresponding to small singular values may not be adequately sensed and the sensor configuration can then be altered to ensure that the system figure and alignment control will meet system requirements.

It should be mentioned that many active optical system designs will fail the test of having no small singular values and still be capable of good performance. One way in which this can occur is that an intentional redundancy is present in the actuator configuration. For example, the same motion of the surface of a mirror panel can be accomplished: (1) by position actuators that move the panel and substrate assembly as a rigid body, and (2) by figure control actuators that connect the mirror faceplate to the substrate. In such a case, there is not one unique setting of the actuators that will properly correct the system. This is not a significant problem in achieving system performance goals. The actuator mode that corresponds to the small singular value in such a case contains the difference between the two actuator setting combinations that produce the same system response. If the singular value is below some threshold, the proper response is to drop that mode from the solution entirely. Dropping modes with zero singular values gives a solution corresponding to the Moore-Penrose pseudo-inverse. The pseudo-inverse has the property that the sum of the squares of the actuator settings associated with the redundancy are minimized. Suitable scaling of the system equations for differing ranges of redundant fine and coarse actuators will result in automatic hierarchical control.

We can systematize the evaluation of optical performance for actuators with significant redundancies by basing the performance criterion on the wavefront averaged over the field of view. If a unit of the ith component of g produces a wavefront error wi, then the expected value for the wavefront error from all modes, W, is approximately given by

$$\langle W^2 \rangle = N^2 \, \mathrm{sum}(w_i^2/s_i^2) \tag{11}$$

Performance prediction

The noise sensitivity of the least squares solution is useful in evaluating alternative sensor and actuator configurations. However, the simple least squares solution often is not adequate for system control because the wavefront sensor noise may be significant with dim stars as sources. The optimum linear estimator for the mode strengths is given by

$$g'_i = \left[s_i/(s_i^2 + L_i^2) \right] (U z')_i \tag{12}$$

where L is set to minimize the error of g' in a least squares sense.

$$L_i^2 = N^2/\langle g_i^2 \rangle \tag{13}$$

The variance of the actuator modes is then given by

$$\langle dg_i^2 \rangle = N^2 \langle g_i^2 \rangle / (s_i^2 \langle g_i^2 \rangle + N^2) \tag{14}$$

Eq. 14 predicts the residual error in the actuator setting modes after the correction is performed. This is the new estimate for actuator mode variance applicable in Eqs. 12 and 13 for subsequent corrections with the addition of the expected variance due to environmental factors. The convergence of the optical figure may be predicted through the evolution implied by successive application of Eqs. 12 through 14.

This approach was applied to a wide angle multipaneled design. The system evolution was predicted using Eq. 14. The systems control was also simulated applying the optimal linear estimation of Eqs. 12 and 13. The evolution of the imaging quality averaged over the field of view is shown in Figure 4. The performance of the same system without edge displacement sensors was also studied. As expected, it was found that actuator setting modes related to piston errors in the panel were not sensed adequately by the wavefront sensors that measure wavefront tilt. The evolution of the system without edge sensors was simulated with the result shown in Figure 5.

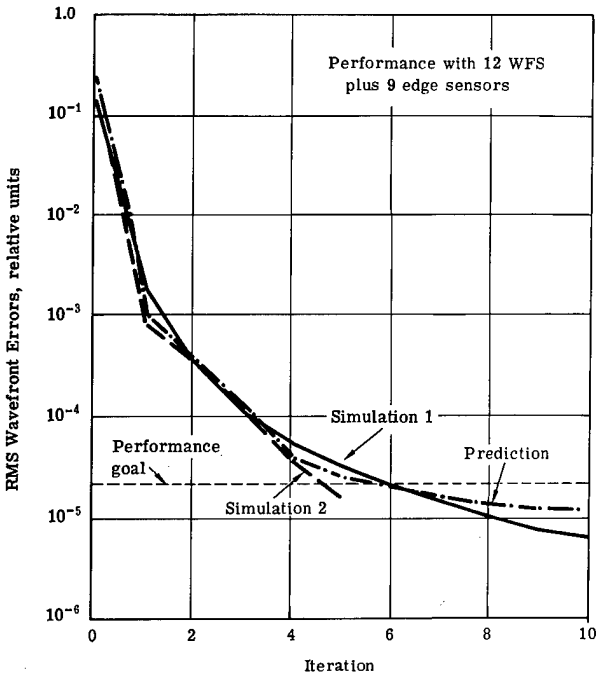


Figure 4. Simulated system control performance using edge phase

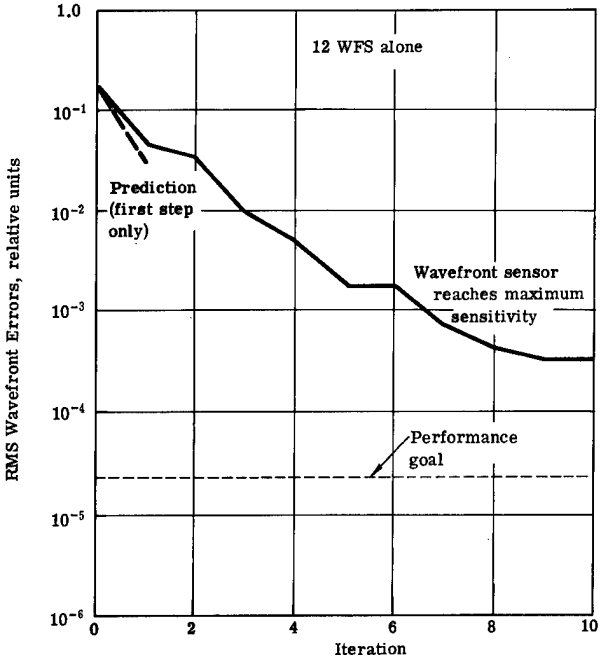


Figure 5. Simulated system control performance

Summary

Very large, multipanel, multisurface optical systems to be built in the future will require active control to provide good performance. The alignment and figuring of such systems will be difficult but it appears to be tractable. We have described our conception of the setup and figuring of these large systems and provided an analysis approach that can be used to determine if the sensor-actuator configuration is adequate for the task and to predict the resultant system performance.

Acknowledgements

The authors wish to thank Paul Noren for the optical model and John Pepi and Joe Lowe for the structural model. We also wish to thank Thomas Pitts for technical guidance.

References

1. Nelson, J., Segmented Mirror for a 10-Meter Telescope, SPIE, 172, Instrumentation in Astronomy III (1979).
2. Wyant, J.C., Use of an AC Heterodyne Lateral Shear Interferometer with Real Time Wavefront Correction Systems, Appl. Opt., 14:2622 (1975).
3. Hardy, J.W., Adaptive Optics: A New Technology for the Control of Light, Proc. IEEE, 66(6) (June 1978).
4. Rimmer, M., Analysis of Perturbed Lens Systems, Appl. Opt., 9:533 (1970).
5. Golub, G.H., and Reinsch, C., Singular Value Decomposition and Least Square Solutions, Numer, Math., 14:403 (1970).
6. Grabow, B.S., Boyle, J.M., Dongarra, J.J., and Moler, C.B., "Matrix Eigensystem Routines — Eispack Guide Extension," Springer-Verlag, Berlin, 1977.
7. Lawson, C.L., and Hanson, R.J., "Solving Least Squares Problems," Prentice Hall, Englewood Cliffs, N.J., 1974.

Adaptive optics without wavefront sensors

Virendra N. Mahajan, Jacques Govignon, Ricky J. Morgan
The Charles Stark Draper Laboratory, Inc., Cambridge, Massachusetts 02139

Abstract

Two sharpness functions are considered for the problem of dynamic phase estimation and correction of aberrated images. One is based on the integral of the square of the image irradiance distribution. The other uses integration of the irradiance distribution over an area smaller than the diffraction-limited resolution area. Examples of phase retrieval using these sharpness functions are given, and laboratory experiments on their verification are described. Applications of this technique to high-energy laser systems, thermal blooming, imaging through atmospheric turbulence, and large optical systems are discussed.

Introduction

An adaptive optical imaging system generally consists of an imaging element, a sensor for determining the wavefront errors of an image forming wave, and an adaptive element (corrector) to compensate for these errors. In this paper we use image sharpness functions[1] for adaptive estimation and correction of aberrations of an incoherent optical imaging system. An aberrated image is corrected by introducing aberrations into an adaptive element until a sharpness function, which is derived from the image irradiance data, is optimized. The sharpness functions have the property that they achieve their extremum values when the image is aberration free. Thus, by forming a closed loop between the image and the adaptive element, an aberration-free image is obtained without the use of any pupil-plane wavefront sensor. This method can also be used for phase retrieval of aberrated images.

Two sharpness functions are considered for the problem of phase estimation and correction of degraded images. One is based on the integral of the square of the image irradiance distribution. The other uses integration of the irradiance distribution over a small area, i.e., it is a measure of the encircled energy of a point-spread function. Examples of phase retrieval using these sharpness functions are given, and laboratory experiments on their verification are described. Applications of this technique to high-energy laser systems, thermal blooming, imaging through atmospheric turbulence, and large optical systems are discussed.

Image sharpness functions

If $I(\vec{r})$ represents the irradiance distribution of an image for a given shape of the pupil and amplitude distribution across it, then the sharpness functions

$$S_1 = \int I^2(\vec{r})\, d\vec{r} \tag{1}$$

and

$$S_2 = \int_{\Delta\vec{r}} I(\vec{r})\, d\vec{r} \tag{2}$$

considered as functions of aberrations are maximum for an aberration-free image. Whereas in Eq. (1) the integration is carried out over the incoherent image of an isoplanatic extended object with a uniform background, in Eq. (2) it is carried out over a small region of the image of a point (i.e., unresolved) object. The sharpness function S_2 is sensitive to a wavefront tilt, but S_1 is not. Therefore, when using S_1, the motion of an image is corrected separately, e.g., by centroid detection and a tilt-control mirror. It is evident that, in practice, the use of S_1 requires sampling of the image by an array of small detectors.

Figure 1 shows how S_1 and S_2 vary, for example, with balanced spherical aberration for a point object. The spherical aberration is balanced with an equal and opposite amount of defocus to yield a minimum variance. The region of integration of the point spread function is a square with a full width of $2\lambda F$, where λ is the optical wavelength and F is the focal ratio (f-number) of the image. It is evident from this figure that when the aberration is large, not only does the sensitivity of the sharpness functions decrease, but also, that they do not necessarily increase with a decrease in the amount of aberration. A similar behavior is found for the other primary aberrations.

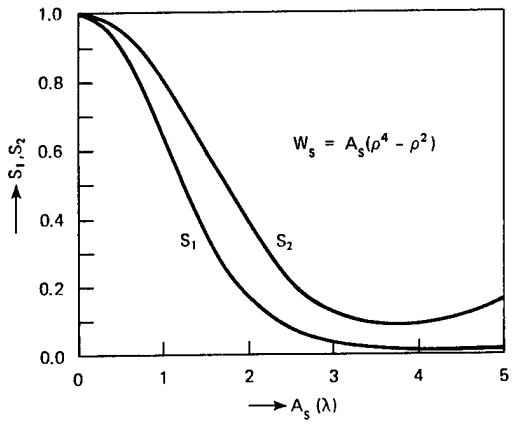


Figure 1. Variation of sharpness functions S_1 and S_2 with balanced spherical aberration.

The problem of sampling the image can be avoided in the case of known extended objects if a mask that has a transmission which is proportional to its unaberrated image irradiance distribution $I_0(\vec{r})$ is placed over a large single detector. Thus, the detector signal given by the sharpness function

$$S_3 = \int I(\vec{r}) I_0(\vec{r})\, d\vec{r} \tag{3}$$

is maximized when the aberrations are zero. Its maximum value is equal to that of S_1. Another possibility is the use of the sharpness function

$$S_4 = \int [I(\vec{r}) - I_0(\vec{r})]^2\, d\vec{r} \tag{4}$$

which has a minimum value of zero when the aberrations are zero. In this case no mask is needed, but again, the image has to be sampled. Both S_3 and S_4 are sensitive to a wavefront tilt.

A comparison of the various sharpness functions is given in Table 1. Note that whenever the image must be sampled, a detector array is used and therefore, more data processing is required. Accordingly, the rate of aberration correction will be lower.

Table 1. Comparison of the Various Sharpness Functions

Sharpness Function	Object	Detector	Bandwidth of Correction	Sensitivity to Tilt
$S_1 = \int I^2\, d\vec{r}$	Point, extended	Array	Low	No
$S_2 = \int_{\Delta\vec{r}} I\, d\vec{r}$	Point	Single	High	Yes
$S_3 = \int I I_0\, d\vec{r}$	Point, known extended	Single	High	Yes
$S_4 = \int (I - I_0)^2\, d\vec{r}$	Point, known extended	Array	Low	Yes

Phase retrieval

Consider an imaging system with a circular pupil of unit radius aberrated by an aberration given by

$$W(\rho,\theta) = A_t\rho\cos\theta + B_t\rho\sin\theta + A_d\rho^2 + A_a\rho^2\cos 2\theta + B_a\rho^2\sin 2\theta + A_c\rho^3\cos\theta + B_c\rho^3\sin\theta + A_s\rho^4 \tag{5}$$

where $0 \leq \rho \leq 1$ and $0 \leq \theta < 2\pi$. Thus the aberration function is made up of tilt (A_t,B_t), defocus A_d, astigmatism (A_a,B_a), coma (A_c,B_c), and spherical aberration A_s. We have considered several combinations of these aberration coefficients such that the peak-to-peak value of the aberration function is $\lesssim 2\lambda$. As an example, let $A_t = -0.376$, $B_t = -0.432$, $A_d = -0.110$, $A_a = 0.354$, $B_a = 0.338$, $A_c = 0.564$, $B_c = 0.648$, and $A_s = 0.636$ where all of the coefficients are in units of the optical wavelength. For simplicity, the amount of coma chosen is such that when combined with the tilt, it gives a minimum aberration variance. A plot of the aberration function and the corresponding point-spread function (PSF) are shown in Figure 2. The peak-to-peak value of the aberration, Strehl ratio, and sharpness value of the image are also given in the figure.

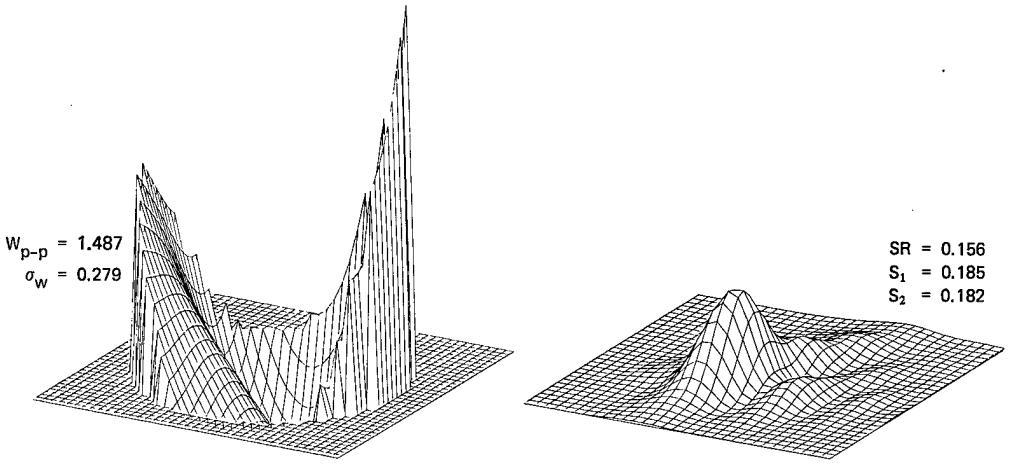


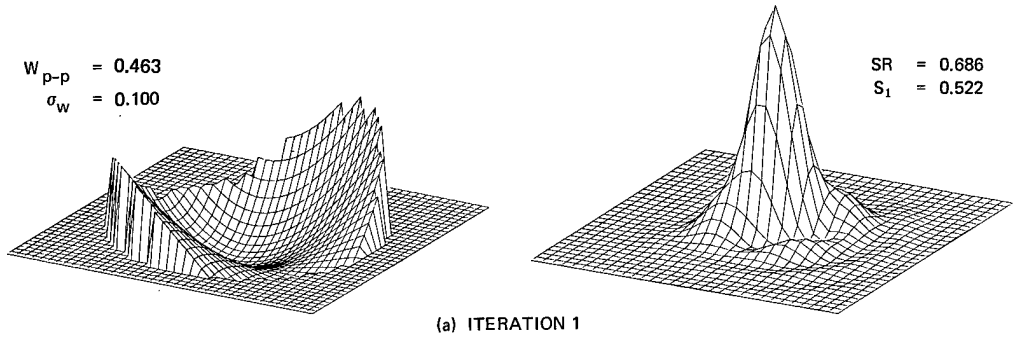
Figure 2. Aberration function and the aberrated image. The peak-to-peak value (W_{p-p}) and the standard deviation (σ_w) of the aberration are given in units of λ. The Strehl ratio (SR) and the image sharpness values S_1 and S_2 are also given.

In our phase retrieval algorithm,[2] aberrations are introduced sequentially in terms of orthonormal Zernike polynomials[3] $\varepsilon_m\sqrt{2(n+1)}R_n^m(\rho)\begin{matrix}\cos m\theta \\ \sin m\theta\end{matrix}$, where $\varepsilon_m = 1/\sqrt{2}$ when $m = 0$ and $\varepsilon_m = 1$ when $m \neq 0$. As a small amount of aberration is introduced, the sharpness function is evaluated, and depending on the change from its previous value, more aberration with positive or negative sign is introduced. From the original and the two new values of the sharpness function, the required amount of correction of that aberration is estimated and a correction is made. This procedure is repeated for each of the polynomial aberrations in the algorithm, and the whole process is repeated iteratively. In practice, the aberrations may be introduced sequentially or in parallel depending upon the bandwidth requirement and affordable complexity. For parallel operation, the aberrations may be introduced with a different harmonic dependence on time for each polynomial aberration, so that their effects on the sharpness function can be discerned.

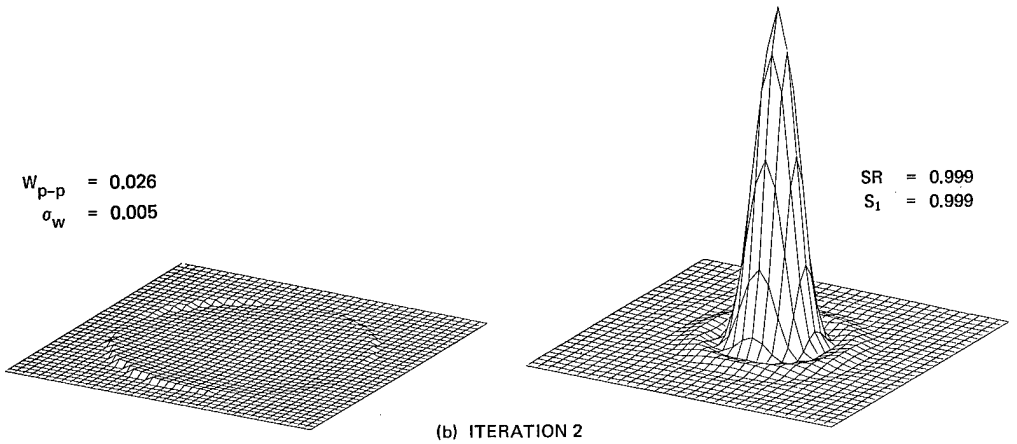
Figures 3 and 4 show the progression of phase retrieval and image correction using sharpness functions S_1 and S_2, respectively. Both the residual aberration function and the corresponding PSFs are shown in these figures. An iteration in these figures is defined to be one pass through all the polynomial aberrations. It is evident from these figures that both sharpness functions lead to an excellent correction of the image in two iterations. The region of integration for calculating S_2 in this example was a square with a full width equal to the radius of the Airy disc.

Experimental verifications

A schematic of the laboratory setup for testing the use of sharpness functions in a closed-loop manner is shown in Figure 5. To simulate a point object, a pinhole is illuminated by a He-Ne laser beam. The light from this object falls on a deformable membrane mirror fabricated by the Perkin-Elmer Corporation.[4] The part of the optical train between the laser and the deformable mirror can also be looked upon as a beam expander. The figure of the membrane is controlled electrostatically by a set of 52 actuators distributed over an area approximately 1 inch in diameter. We wired the actuators in such a way that the membrane figure can be controlled in terms of Zernike modes. When the controller is off,

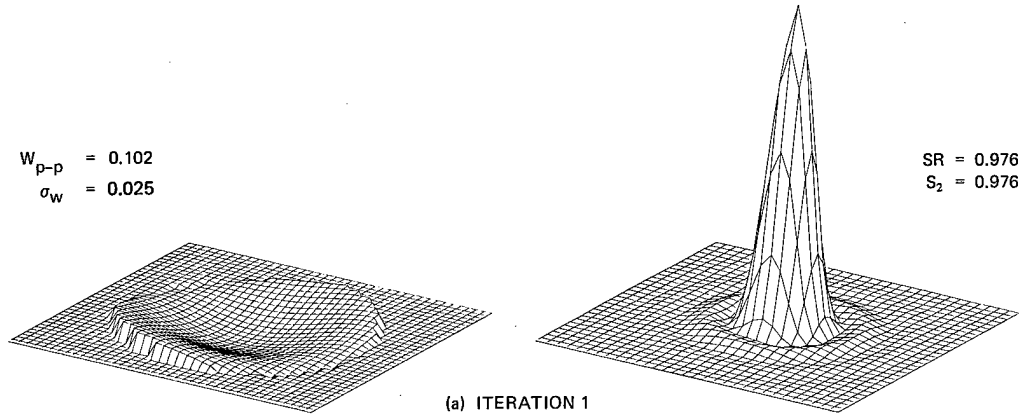


(a) ITERATION 1

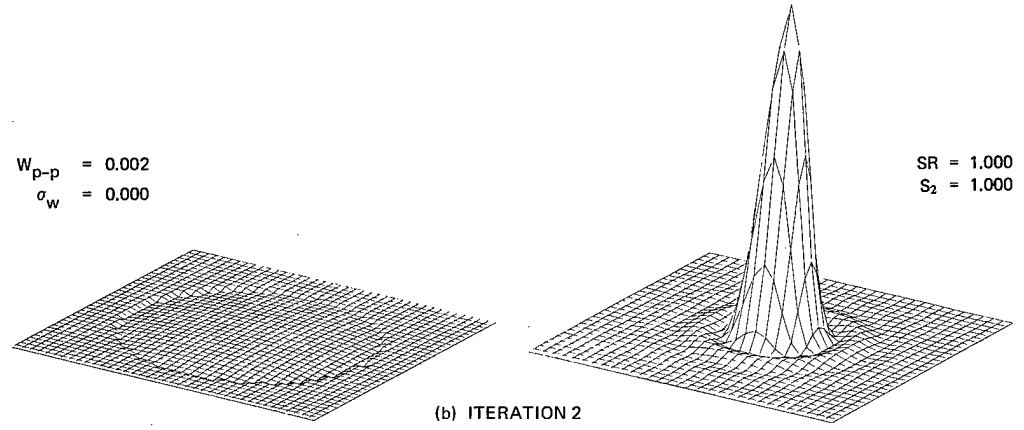


(b) ITERATION 2

Figure 3. Phase retrieval using image sharpness function S_1. The residual wavefront errors and the corresponding images after each iteration are shown.



(a) ITERATION 1



(b) ITERATION 2

Figure 4. Phase retrieval using image sharpness function S_2.

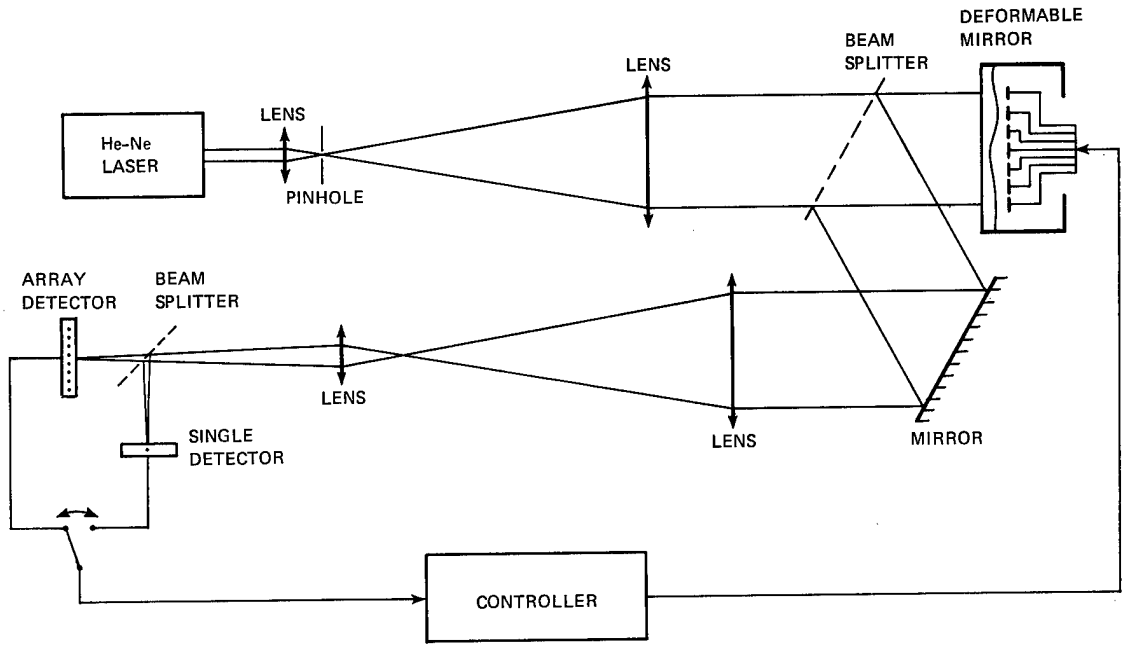


Figure 5. Schematic of the laboratory setup for adaptive correction of aberrated images.

the figure is such that it gives the interference fringes shown in Figure 6a. These fringes also include any errors in the rest of the optical train. The f-number of the imaging optics, which has a circular pupil, is approximately 600. The image as observed by a 32 × 32 photodiode array (supplied by Reticon Corporation), with center-to-center spacing of 100 μm between two adjacent photodiodes, is shown in Figure 6b. This array with its associated processor was also used to measure the sharpness function S_1. An appropriately sized single detector was used to measure the value of the sharpness function S_2.[5] A digital controller addresses the actuators to deform the mirror in various Zernike modes. An example of a corrected image obtained by addressing six Zernike modes is shown in Figure 6c. These modes consisted of defocus (n = 2, m = 0), two astigmatisms (n = 2, m = 2), two comas (n = 3, m = 1) and spherical aberration (n = 4, m = 0). No observable difference between the quality of the images corrected by using sharpness functions S_1 and S_2 was obtained.

(a) ABERRATED WAVEFRONT (b) ABERRATED IMAGE (c) CORRECTED IMAGE

Figure 6. Adaptive correction of an aberrated image using sharpening.

Applications

The image sharpening technique described here is applicable to many problems where adaptive optics is needed. These include deformable high-energy laser (HEL) systems, thermal blooming, imaging through turbulent media, and disturbed large optical systems. Some of the HEL systems require adaptive optics due to all of these factors. The optical elements (e.g., mirrors) of an HEL system distort due to the high-power densities of the laser beam. We have shown[6] that for circularly symmetric beams and an on-axis target, most of the thermally induced distortion is defocus with a small amount of spherical aberration. For an asymmetric beam and an off-axis target, tilt, coma, and astigmatism will also be introduced. Thus, for the thermally introduced aberrations of the system, correction of a few Zernike modes will be adequate.

For an HEL beam propagating through the atmosphere, depending on its power density, the refractive index of the atmosphere can change significantly and thus alter the phase distribution across it. This phenomenon, called thermal blooming, can severely degrade the target-plane irradiance distribution. However, by preaberrating the beam (e.g., with a deformable mirror), the phase errors can be compensated in many situations.

The atmospheric turbulence can not only degrade the target-plane irradiance distribution, but it can also degrade the image of the target observed at the HEL system. In fact, atmospheric turbulence can degrade any imagery encountering it. The time-averaged variance of aberrations introduced by Kolmogoroff atmospheric turbulence is given by[7]

$$\sigma_t^2 = 0.0261(D/r_0)^{5/3}\lambda^2 \tag{6}$$

where D is the pupil (beam) diameter, and r_0 is a characteristic length of the atmospheric turbulence.[8] If the wave aberration caused by atmospheric turbulence is decomposed into an infinite set of Zernike modes, the reduction in variance of the aberration after correction of the first J modes is given by

$$\sigma_c^2 = \sum_{j=2}^{J} \langle a_j^2 \rangle \tag{7}$$

where, the time-averaged variance of a modal aberration is given by

$$\langle a_j^2 \rangle = 0.0192 \frac{(n+1)\Gamma(n-5/6)}{\Gamma(n+23/6)} (D/R_0)^{5/3}\lambda^2, \quad j \geq 2 \tag{8}$$

In Eq. (8), $\Gamma(\cdot)$, is the gamma function. The mode with j = 1 represents a uniform phase error across the pupil and therefore does not contribute to the aberration variance. The relationship between the indices n, m, and j is given in Table 2. Note that the variance of a modal aberration with a given value of n depends on n but not on m. Thus, all modes that have the same value of n but different values of m contribute equally to the aberration variance.

Table 2. Relationship between indices n, m, and j. (Note that for every nonzero value of m, there are two j values corresponding to the modes with cos θ and sin θ dependence.

n	m	j
0	0	1
1	1	2,3
2	0,2	4,5,6
3	1,3	7,8;9,10
4	0,2,4	11;12,13;14,15
5	1,3,5	16,17;18,19;20,21
6	0,2,4,6	22;23,24;25,26;27,28

The standard deviation of the residual wave aberration after correction of the first J modes is given by

$$\sigma_J = (\sigma_t^2 - \sigma_c^2)^{1/2} \tag{9}$$

Figure 7 shows how σ_J varies J. It is clear that the first few modes contain a large fraction of the error. The reduction in error per mode decreases as the mode number increases. Thus, in situations were correction of a small number of modes gives an acceptable level of correction, the modal technique is very attractive. For example, if $\sigma_J \lesssim \lambda/10$ is acceptable, then for $D/r_0 \lesssim 5$, $J \lesssim 15$. The first 15 Zernike modes include all the aberrations up to the fourth order.

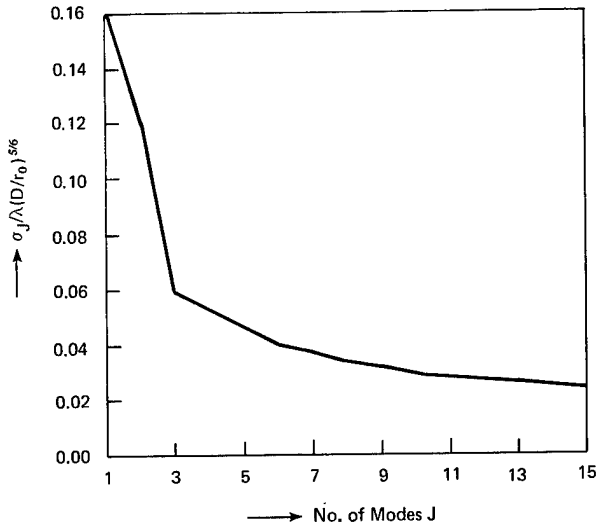


Figure 7. Variation of σ_J with J. (Note that the first mode, which represents uniform phase error across the pupil, does not affect the aberration variance.)

Large optical systems, especially the ones for space applications, use lightweight mirrors to reduce the weight and material cost. These lightweight mirrors are flexible and therefore deform when the system is disturbed. However, again most of their distortion is made up of the first few Zernike modes. Thus, the modal adaptive optics described herein are applicable to many situations of practical interest.

Acknowledgment

The work reported herein was performed under internal funding at The Charles Stark Draper Laboratory, Inc. The authors gratefully acknowledge the help and encouragement of personnel at the Rome Air Development Center. The hardware for the experimental work was developed under the direction of W. P. Curtiss of The Charles Stark Draper Laboratory, Inc.

References

1. Muller, R.A., and A. Buffington, "Real-Time Correction of Atmospherically Degraded Telescope Images Through Image Sharpening", J. Opt. Soc. Am., Vol. 64, pp. 1200-1210, 1974.

2. Welling, P.A., Optimization of Image Sharpness in an Adaptive Optical System, Thesis, Massachusetts Institute of Technology, 1978. This reference contains early work on the development of the algorithm.

3. Born, M., and E. Wolf, Principles of Optics, 5th ed., (Pergamon, New York, p. 464, 1975.

4. Grosso, R.P., and M. Yellin, "The Membrane Mirror as an Adaptive Element," J. Opt. Soc. Am., Vol. 67, pp. 399-406, 1977. See also, M. Yellin, "Using Membrane Mirrors in Adaptive Optics", Proc. SPIE, Vol. 75, pp. 97-102, 1976.

5. Buffington, A., F.S. Crawford, R.A. Muller, A.J. Schwemin, and R.G. Smits, "Correction of Atmospheric Distortion with an Image-Sharpening Telescope", J. Opt. Soc. Am., Vol. 67, pp. 298-303, 1977. A one-dimensional analogue of our experiment is described in this reference.

6. Mahajan, V.N., and J. Govignon, "Computer Simulation of a Large Adaptive Optical System", Proc. SPIE, Vol. 72, pp. 439-451, 1979.

7. Noll, R.J., "Zernike Polynomials and Atmospheric Turbulence", J. Opt. Soc. Am., Vol. 66, pp. 207-211, 1976.

8. Fried, D.L., "Statistics of a Geometric Representation of Wavefront Distortion", J. Opt. Soc. Am., Vol. 55, pp. 1427-1435, 1965.

Absolute distance interferometry

C. W. Gillard, N. E. Buholz, D. W. Ridder
Sensor Technology Department, Lockheed Missiles and Space Company
O/62-12, B/151, P.O. Box 504, Sunnyvale, California 94088

Abstract

A basis for making absolute distance measurements to an accuracy of 0.025 μm over 0-1.5 m intervals is reported. Extensions of this technology will permit distances of 50 m and greater to be measured to the same accuracy. Two-color, synthetic Michelson interferometry using a CO_2 laser source capable of generating four sets of R- and P-line pairs is employed. This allows reduction of the very large ambiguity exhibited by conventional Michelson interferometers as well as the resolution of difficulties which would otherwise arise due to instabilities in the measurement arm of the interferometer. This latter effect is a practical rather than a fundamental consideration, but is nonetheless important if the interferometer is to emerge from the laboratory as an effective, workable instrument. Distance is determined in terms of a denumerable number of precisely known wavelengths and fractions thereof.

Introduction

The measurement of distance in terms of optical wavelengths over extended ranges was first achieved late in the last century. Undoubtedly the most famous such measurement was the determination of the standard meter in terms of the monochromatic red cadmium line by Michelson and Benoit over eight decades ago. A desire for making precise, absolute measurements over extended distances in terms of optical wavelengths thus has been recognized for many years; but making such measurements was a feat of virtuosity, a real tour de force. Turning this desire into a practical reality was to await the advent of technology but recently available: primarily, the laser and the small, dedicated computer.

Basis of absolute distance interferometry

Anyone pondering the problem of making absolute distance measurements with a Michelson interferometer soon discovers a need to contemplate the principles of obtaining and employing fractional fringe measurements, of multi-color interferometry, and of generating multiple colors of the requisite spacing through the use of stabilized, multi-color lasers. In searching prior work covering these areas, only two papers have been found which report on work of a directly related nature: one by Tilford[1], and the other by Bourdet and Orszag[2] in France. It seems that despite the apparent rewards in developing a practical means of making accurate, absolute distance measurements, the technical challenges are difficult, the expense considerable, and the rewards, in terms of applications of substance, may be slow in materializing..Thus, it should not be too surprising to find the field uncrowded.

Tilford develops a mathematical logic and points out the feasibility of employing a multifrequency, stabilized CO_2 laser source and fractional fringe measurements for the purpose of making absolute measurement of distance. Tilford's work is limited to that of a theoretical analysis. Bourdet and Orszag present experimental data. They employed a grating-tuned CO_2 laser as a multi-color source, a Michelson interferometer, and measured fractional fringes by tuning the reference arm of the interferometer and measuring the displacement. They achieved an accuracy of 0.1μm where the distance to be measured was known a priori within 1 m. Difficulties arising from the lack of stability of the reference arm of the Michelson interferometer were experienced. Both of these papers provide pertinent background for the developments reported herein.

It should be mentioned that Bourdet and Orszag's paper appeared after the work reported herein was well underway, and that the two approaches are thus independent. But despite their independent origins, the basic philosophies are amazingly similar. The differences in approach which do exist, however, are more than procedural: they are substantive and even crucial to applications in areas where speed of measurement is important and a laboratory environment is out of the question.

The purpose of this paper is to present an overview of the approach, discuss the basic elements, and review the progress and results achieved to date of an experimental system designed to measure absolute distance interferometrically. Papers on the various subsystems, as well as operational processes, will appear as the work evolves.

Approach

In addressing the goal of absolute distance interferometry, a number of decisions were made at the outset which, in turn, circumscribed the manner in which the various subsystems were implemented. A review of the decisions and of basic approaches to the major subsystems of the absolute distance interferometer follow; included are brief comments regarding their selection.

Heterodyne photodetection

Heterodyne photodetection provides a basis for quantum-noise-limited operation, gives high S/N ratios and a large signal with small target return power. It also accommodates a large variation in signal level (wide dynamic range) and permits the phases (fractional fringes) of the optical signals to be measured electronically at a convenient frequency. Optical offsets necessary for heterodyne operation are provided by Bragg cells. The cells also serve to spectrally isolate the laser from target returns which would otherwise upset its stability.

Laser

A one meter CO_2 laser, stabilized and designed to operate simultaneously on several R- and P-line wavelength sets, is employed as source. Its selection was based on what appears to be a unique set of characteristics: the CO_2 laser is the only known laser with the abundance of

appropriately spaced, well characterized wavelengths capable of meeting the fractional fringe criteria employed to determine length. Other lasers may exist which are amenable to these criteria, but the CO_2 laser appeared to be the only one suitable. Simultaneous R- and P-line operation provides the basis of two very important features which will be discussed in detail later: (1) it provides absolute control over the CO_2 laser wavelengths; (2) by creating differential fringes derived from simultaneous R- and P-line phase measurements, the distance measurement process becomes insensitive to instabilities in both the reference as well as the target arms of the synthetic Michelson interferometer. It should also be mentioned that with the CO_2 laser, high power is available (10 W per line for a 1 m laser), a fact which opens up the possibility of ranging against unprepared, nonspecular targets.

Interferometer

The breadboard interferometer (Figure 1) is not a true Michelson interferometer: it may be considered a synthetic Michelson interferometer with a subordinate heterodyne interferometer. The heterodyne interference pattern is actually a dynamic event, periodic at the heterodyne frequency, where the "reference" leg of the heterodyne interferometer is made up of both the optical local oscillator (LO) beam, starting from the point that LO beam separation occurs, and an electronic signal phase path which is related to the local oscillator offset frequency. Note that the target and reference arms of the synthetic Michelson interferometer are viewed separately and sequentially. Phase subtraction is thus an electronic rather than an optical process. By measuring the reference and target phases (distances) separately and then subtracting, variations due to instabilities in the electronics as well as the optics upstream from the synthetic Michelson interferometer beamsplitter become common to both the target and reference paths, and are largely removed by the subtraction process.

Experimental apparatus

For convenient reference we shall refer to the measurement system described herein as the "absolute distance sensor" (ADS). Elements of the ADS system are briefly described next. A review of some of the significant results achieved to date is discussed in the following section.

ADS breadboard

Details essential to a basic understanding of the optical configuration of the ADS breadboard are illustrated in Figure 1. Starting with the CO_2 laser, it is noted that the laser emits power from both ends. Power from the upper end in the figure is used to control and stabilize the laser. Details of laser control and stabilization will be discussed later. Power from the lower end of the laser is split into two beams, one of which is employed in the interferometer, and the other as local oscillator. As previously mentioned, heterodyne photodetection is used, which implies that the LO beam, interferometer beam, or both, are frequency offset. In the case of the ADS breadboard, both beams are frequency offset. Bragg cells provide the means of frequency translation. Ordinarily a single frequency translation would suffice; but a single Bragg cell generates, in addition to the desired frequency-translated component, a second contaminating component due to a small backward (reflected) acoustical wave. This component is large enough to impair the ADS phase measurement process. Employing a second Bragg cell at a second frequency provides a basis for isolating this component.

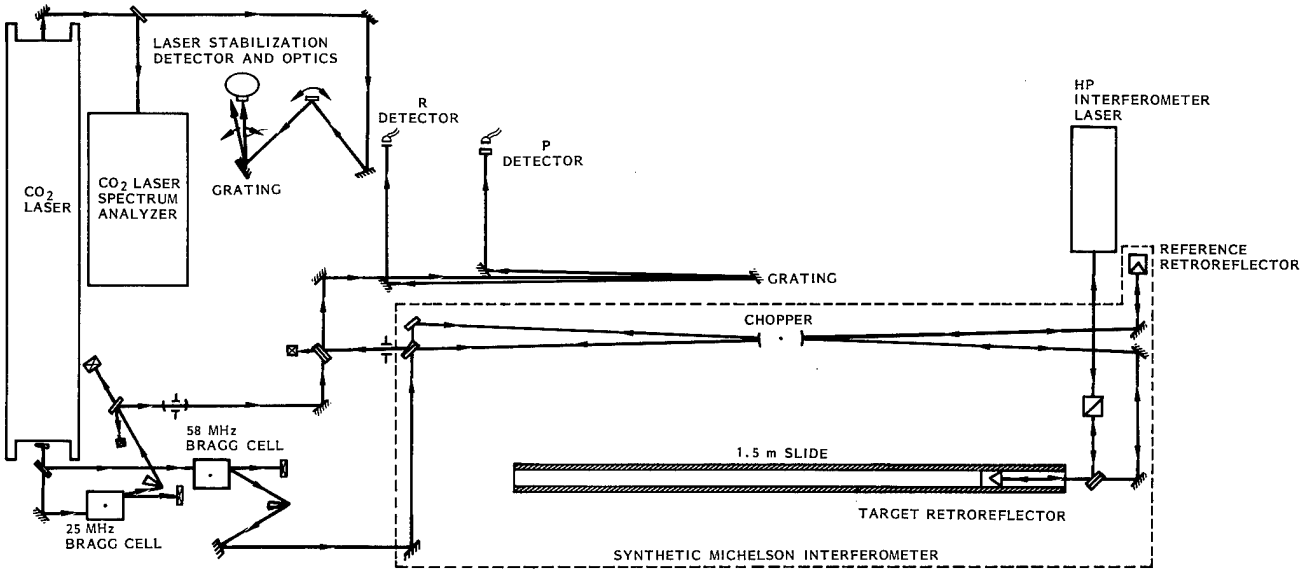


Figure 1. Absolute distance sensor breadboard schematic

Progressing through the optical train, separate target/reference sensing beams are established at a beamsplitter and pass in the vicinity of an optical switch (chopper) employed to alternately range to the target and reference legs of the "interferometer". Since it is the difference between these measurements that is of interest, phase noise in the electronics and all optics upstream from the beamsplitter is common to both measurements and may be largely cancelled. By making rapid phase measurements and by switching between target and reference at a high rate (240 Hz is used here) significant common mode noise cancellation benefits may be achieved. The target/reference beams leaving the interferometer are made colinear at the beamsplitter, are made congruent with the LO beam at a second beamsplitter, and progress to a grating where the R- and P-lines are separated and directed to separate detectors. The heterodyne photodetector signals are then digitally processed for phase as illustrated in Figure 2.

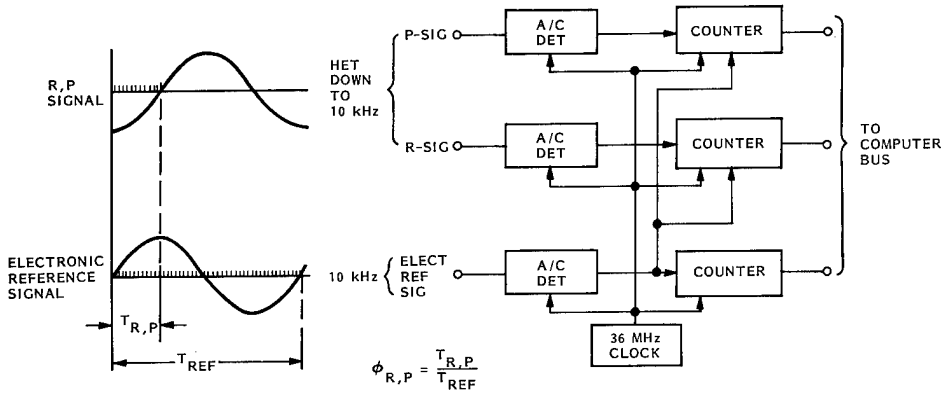


Figure 2. ADS digital phase meter

In the scheme illustrated, a 36 MHz clock is started and stopped by axis crossing detectors using target and reference returns from the detected R- and P-carriers as well as an electronic reference signal derived from the Bragg cell drivers, all of which are heterodyned to a convenient working frequency (in this case, 10 kHz). Statistically, the digital error of this scheme is approximately 1:1.8 x 10^4 for each phase measurement. Statistical improvement is achieved by averaging over 10 cycles per phase measurement. Relative times (reference signal period and target return delay times) are directly available in a form convenient for computer input without further processing. Although inefficient in terms of complete utilization of the available signal, this phase meter is fast, simple, flexible and reliable.

Two-color laser

The key element of ADS is the "two-color laser" (T-C laser), which is capable of stabilizing and operating simultaneously on any of four sets of two-color pairs (for a total of five different rotational-vibrational lines in the 10.4μm band), and of rapidly switching through the various color pairs by means of a piezoelectric mirror drive and control subsystem. All of the basic features of the T-C laser (its stability, switching capability, line pairing sequence, states-of-operation, switching speed, and derivable wavelength hierarchy) are fundamental to the operation of the ADS system.

The basic concept behind the T-C laser stabilization technique is related to the fact that a linear power exchange exists between the R- and P-lines as a function of laser frequency in the region of two-color operation. Equalizing the R- and P-line powers thus provides a laser frequency discriminant which is based on the CO_2 molecule and which, in principle, should be absolute.[3] Laser stabilization test data obtained by beating two stabilized, nearly identical T-C lasers together indicates that the residual noise of one T-C laser is approximately ±35 kHz, or 1 part in 0.8 x 10^9/15 minutes. Short term stability is 1 part in 7 x 10^9/50 msec. Figure 3 schematically displays the basic laser stabilization arrangement used in the ADS system. Papers dealing with specific areas of the four-state T-C laser and stabilization scheme will be published in the near future.

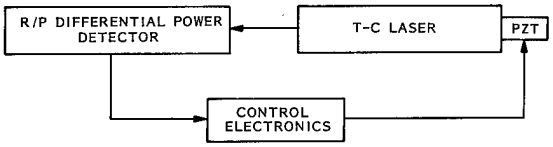


Figure 3. T-C laser stabilization circuit

The signature of the T-C laser (dominant resonant frequency as a function of laser cavity length) is chosen to give the desired sequence of R-, P-lines in the 10.4μm vibrational band. A piezoelectric mirror mount in conjunction with the stabilization circuit adjusts the cavity length to correspond to a chosen segment of the signature. Adjusting the cavity length over a portion of a specific 5μm range allows one to obtain the signature segment shown in Figure 4. Also shown here are the four color-pairs, the off-line-center stabilized operating points, and the four-state switching sequence. Figure 5 illustrates the control electronics input signal as the laser is cycled through the switching sequence. "Servo locked" indicates the region of stabilized T-C operation.

Another way of viewing the five laser lines and their pairing, as well as the deviations from line center, is shown in Figure 6. Here the line-pairing at the bottom corresponds to the states-of-operation in Figure 4. The wavelength and frequency values shown are the line-center values, known to within 100 kHz.[4] The operating frequencies are actually offset from line center by the amounts shown, so that a total of eight different frequencies are utilized. The offsets shown represent the results of calibration studies involving the entire ADS system.

Figure 7 displays the hierarchy of differential and synthetic frequency lines obtained from the eight optical frequencies shown on the first (bottom most) level. The second level consists of differential (beat) frequencies of the various two-color line pairs (simultaneous data). The third, fourth, and fifth levels are "synthetic" frequencies since they are derived using more than one color pair (non-simultaneous data) and hence are not physically observable frequencies. The sixth and seventh levels are also "synthetic" since they rely upon data obtained from either side of line center (two different color pairs). The wavelength values are calculated from the frequency values using the speed of light, c = 2.997925(00) x 10^8 m/sec, and assuming a refractive index, n = 1.

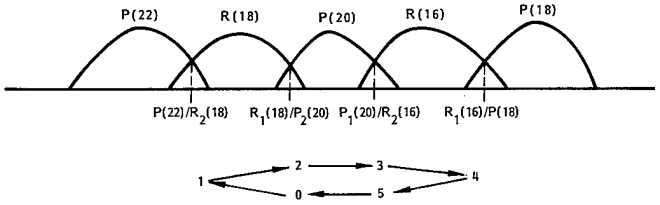


Figure 4. Switching sequence schematic of four state T-C laser

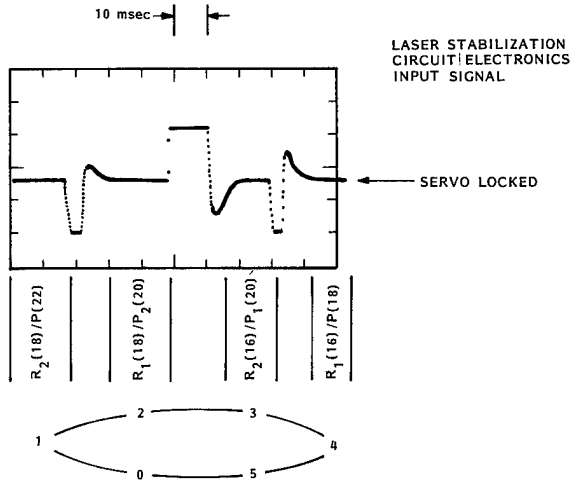


Figure 5. Transient switching properties of the four-state T-C laser

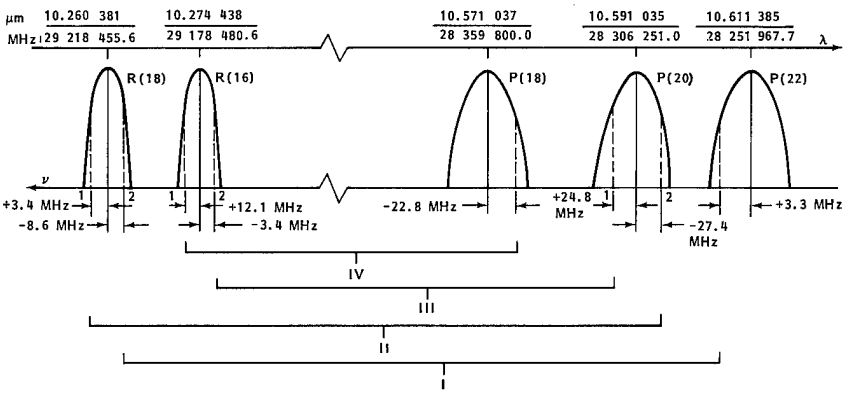


Figure 6. Line pairing schematic of the four-state T-C laser

The ADS system requires two modifications to the frequency and wavelength data presented in Figure 7, one for the refractive index of air and one for the Bragg cell. The refractive index of air is a complex function of wavelength, composition, temperature, pressure, and humidity. A value for the refractive index was calculated using widely accepted formulas for dispersion and atmospheric conditions.[5,6] Table 1 presents wavelength values adjusted for air.* In addition, the laser line frequency values have been adjusted for a 58 MHz downshift due to passage through the first Bragg cell (refer to Figure 1).

A total of sixteen different paths can be traced through the hierarchy pyramid (Figure 7). Four of these paths are presently used in the ADS system, both as a diagnostic for the system and as a means of obtaining information on all eight laser line frequencies. Table 2 displays these four paths using information from Table 1. Eventually, only Path 2 will be used, thus reducing the computer program complexity and increasing computational speed.

Derivation of range

A digital phase meter (described earlier) determines phase information from the heterodyne detected signals of the R- and P-lines. This phase information is then processed to yield range measurements using a dedicated Cromemco Z2D microcomputer operating at 4 MHz. The microcomputer controls laser switching, phase data collection, display of results, and system calibration. However, the rapid conversion of raw phase data into range with values good to 2 parts in 10^8 per meter is the primary function of the microcomputer.

Phase information from each of the four color pairs (for both the target and reference range) is determined by the phase meter and stored in the computer as a fractional fringe number. From this phase information, fractional fringe values are then calculated for each successively higher level of the hierarchy pyramid. Having obtained a complete hierarchy of fractional fringes, range determination may be viewed as a process of successive approximation. The range is first determined at the longest wavelength (hierarchy Level 7) by determining the integral number of the whole fringes using fractional fringe data and an initial range estimate. This new range value is then used, along with the fractional fringe data from the next lower level, to determine a second range value. This process is repeated down through the hierarchy until a final range value is determined at the 10μm wavelengths.

*The atmosphere was nominally maintained near 22°C, 760 torr, and 50% relative humidity, though small fluctuations were usually apparent.

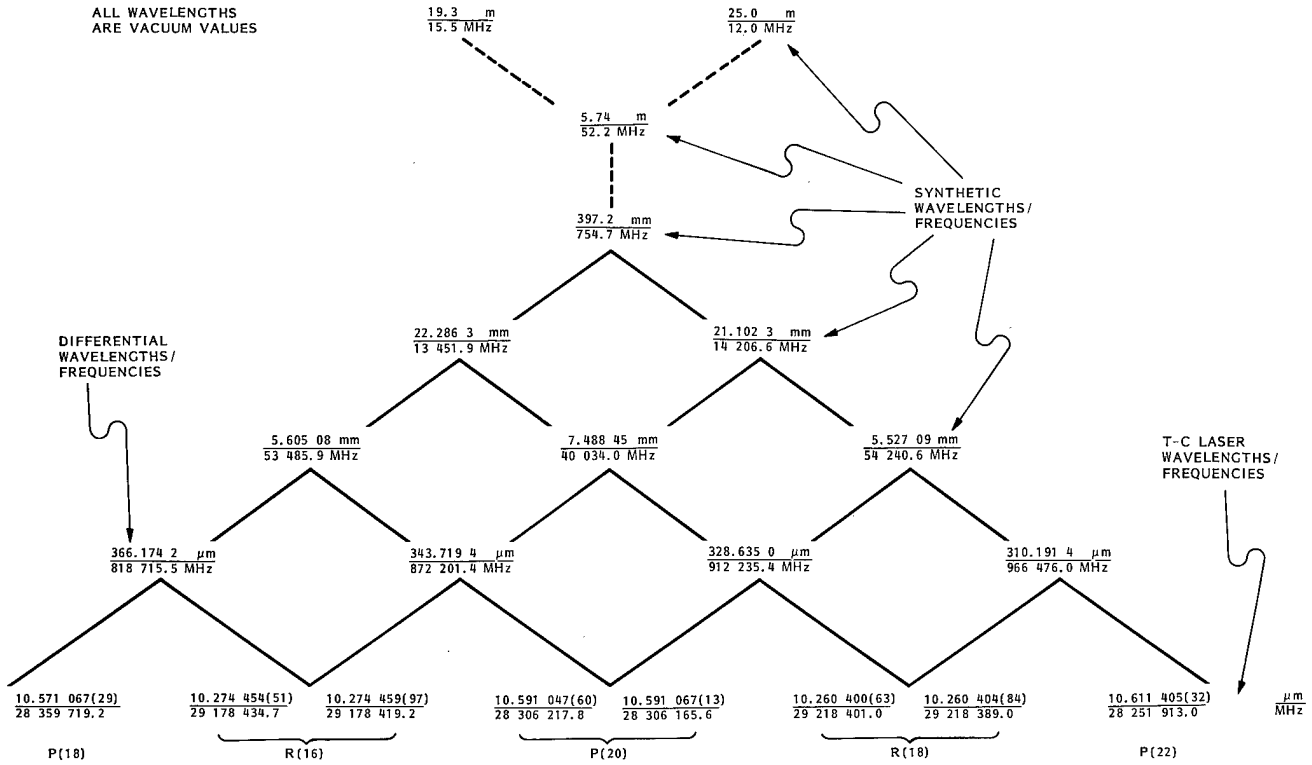


Figure 7. Four-state T-C laser frequency/wavelength hierarchy pyramid

Table 1. Four-State Wavelength Hierarchy (corrected for air)

Frequency MHz	Wavelength	Determination of Frequency and Fringe Values
12.0	25.0 m	$R_1(18) - R_2(18) = \Delta R(18)$
15.5	19.3 m	$R_1(16) - R_2(16) = \Delta R(16)$
52.2	5.74 m	$P_1(20) - P_2(20) = \Delta P(20)$
754.7	397.1 mm	$P_2(20) - P(22) - [P(18) - P_1(20)] - \Delta R(18) + \Delta R(16)$
14206.6	21.0967 mm	$P_2(20) - P(22) - [R_1(18) - R_2(16)] - \Delta R(18) - \Delta P(20)$
13451.9	22.2803 mm	$P(18) - P_1(20) - [R_1(18) - R_2(16)] - \Delta R(16) - \Delta P(20)$
54240.6	5.52562 mm	$P_2(20) - P(22) - \Delta R(18)$
40034.0	7.48646 mm	$R_1(18) - R_2(16) + \Delta P(20)$
53485.9	5.60359 mm	$P(18) - P_1(20) - \Delta R(16)$
966476.0	310.1090 μm	$R_2(18) - P(22)$
912235.4	328.5478 μm	$R_1(18) - P_2(20)$
872201.4	343.6281 μm	$R_2(16) - P_1(20)$
818715.5	366.0770 μm	$R_1(16) - P(18)$
29218389.0	10.257680(67) μm	$R_2(18)$
29218401.0	10.257676(46) μm	$R_1(18)$
29178419.2	10.271732(07) μm	$R_2(16)$
29178434.7	10.271726(61) μm	$R_1(16)$
28251913.0	10.608587(97) μm	$P(22)$
28306165.6	10.588255(18) μm	$P_2(20)$
28306217.8	10.588235(66) μm	$P_1(20)$
28359719.2	10.568260(65) μm	$P(18)$

Table 2. Wavelength Hierarchy

Hierarchy Level	Path #1	Path #2	Path #3	Path #4
7	25 m	25 m	19 m	19 m
6	5.7 m	5.7 m	5.7 m	5.7 m
5	397 mm	397 mm	397 mm	397 mm
4	21.1 mm	21.1 mm	22.3 mm	22.3 mm
3	5.53 mm	7.49 mm	7.49 mm	5.60 mm
2	310 μm	329 μm	344 μm	366 μm
1	R_2 (18) P (22)	R_1 (18) P_2 (20)	R_2 (16) P_1 (20)	R_1 (16) P (18)

The final range value is good to 0.025μm as long as the measurement accuracy (phase resolution) and the laser frequency stability are sufficient to correctly determine the integral number of whole fringes for each wavelength down through the hierarchy. Each wavelength transition in the hierarchy places different numerical criteria for measurement accuracy and laser stability, as seen in Table 3, on the system. These criteria were derived by assuming a total acceptable error of $\lambda/10$ for each wavelength, λ_n, in the hierarchy.

Table 3. Summary of the T-C Laser Frequency Stability and Measurement Accuracy Requirements

Hierarchy Level Transition	Required Measurement Accuracy	Required Laser* Frequency Stability
λ_7: 25 m $\rightarrow \lambda_6$: 5.7 m	$\leq 23.0 \times 10^{-3}\ \lambda_7$	$\leq$ 2.15/L MHz
λ_6: 5.7 m $\rightarrow \lambda_5$: 0.4 m	$\leq 7.0 \times 10^{-3}\ \lambda_6$	$\leq$ 0.74/L MHz
λ_5: 400 mm $\rightarrow \lambda_4$: 21 mm	$\leq 5.3 \times 10^{-3}\ \lambda_5$	$\leq$12.3/L MHz
λ_4: 21 mm $\rightarrow \lambda_3$: 7.5 mm	$\leq 25.0 \times 10^{-3}\ \lambda_4$	$\leq$28.5/L MHz
λ_3: 7.5 mm $\rightarrow \lambda_2$: 0.33 mm	$\leq 4.4 \times 10^{-3}\ \lambda_3$	$\leq$15.1/L MHz
λ_2: 330 μm $\rightarrow \lambda_1$: 10.4 μm	$\leq 3.1 \times 10^{-3}\ \lambda_2$	**
λ_1: 10.4 μm $\rightarrow$ 0.025 μm	$\leq 4.7 \times 10^{-3}\ \lambda_1$	$\leq$ 0.71/L MHz

*L is range measured in meters.
**No stability requirements for simultaneous measurements.

The requirements set in Table 3 are intentionally conservative (by an order-of-magnitude). Even so, the phase meter provides a measurement accuracy of $\pm 1.4 \times 10^{-3}$ fringe ($\pm$0.05 degrees) which is double the most stringent measurement accuracy requirement at the λ_2 level in the hierarchy transition ($\pm 3.1 \times 10^{-3}$ fringe). With range L set at 5 meters, an RMS laser stability of $\pm$140 kHz is required. This, for contrast, is to be compared with the measured stability of the laboratory T-C laser, which is $\pm$35 kHz.

The analysis up to this point has neglected any changes in optical path length which occur during the collection of phase data. Small vibratory excursions do occur which result in inconsistent phase measurements. Phase measurements with a spread of greater than 3% may prevent accurate determination of the range. The role of simultaneous R- and P-line phase measurements in the reduction of system sensitivity to path length changes has been analyzed. Results indicate a significant increase in synthetic wavelength phase measurement accuracy using simultaneous measurements. This increase in accuracy is proportional to the ratio of wavelengths between Level 1 and Level 2 of the hierarchy pyramid (approximately 35:1). The simultaneous measurement of color-pairs provides the ADS system with a crucial insensitivity to path length changes due to target vibration and atmospheric turbulence.

Experimental results

To experimentally obtain a measure of ADS accuracy as well as determine its stability and general performance properties, comparison measurements were made using an HP Interferometer (Model 5525) as reference. The general configuration of this arrangement is illustrated in the breadboard schematic in Figure 1. In an effort to establish coincident measurement geometries the ADS and HP beams were accurately colocated, a common retro-reflector target was employed, and the target-to-reference distances were set nearly equal for the two systems.

Results obtained with this breadboard setup are illustrated in Figure 8. Note that the ADS-HP differential distance measurements vary nearly linearly with range, and that the statistical variation from linearity over the 0-1.5 m target excursion is within $\pm$0.025μm. Experience with the breadboard shows that these results are typical provided the temperature in the vicinity of the ADS "interferometer" is held constant to within approximately $\pm$0.01°C during the measurement process (a few minutes).

Perfect agreement between ADS and HP would yield measurement data all in a straight line with zero slope. The data shown in Figure 8, however, show good linear relationship between ADS and HP, but an apparent wavelength discrepancy exists, indicated for the most part, by the linear bias of 1.30μm/m. The "linear bias" represents an approximate straight-line best-fit to the ADS-HP differential measurement data. This linear portion is removed in the computer and the residuals plotted, using an expanded scale to more clearly reveal statistical variations. Source of linear bias error, which could stem from several areas, is believed to be due to an error in the formula used to correct the CO_2 wavelength for the atmosphere. Value of the calculated atmospheric correction is -265.57μm/m. A 0.5% error in this correction could explain the "linear bias" discrepancy.

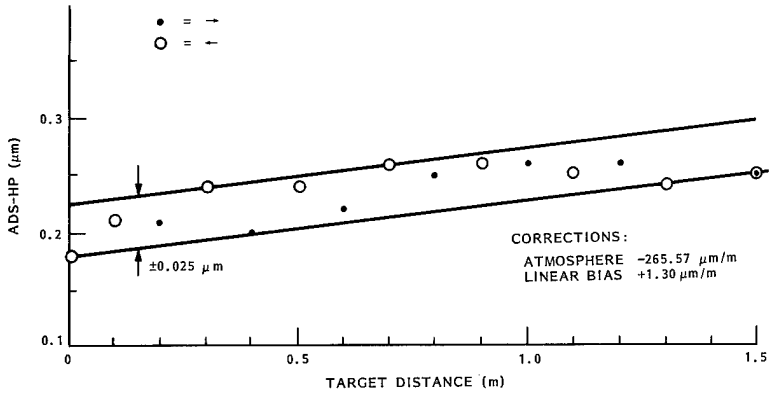


Figure 8. ADS-HP comparison measurements

Applications

The scope of possible ADS applications is very large. To date, however, its practicality has only been demonstrated, and its state-of-development is limited to that of a laboratory breadboard device. Exposure to areas of potential application, both current and future, has been minimal.

An ability to make precise, linear absolute distance measurements represents a useful capability, but the real pay-off to such a capability results when it is employed in conjunction with a beam scanning adjunct. ADS plus a beam scanner provides the basis for accurate measurement of three dimensional structures, and thus, as an example, the means of assisting with the deployment and control of large-scale adaptive optical systems.

Figure 9 presents in schematic form the essential features of a conceptual beam scanning system for ADS (refractive optics would not be employed). It is designed to compensate the beam scanning operation for the gross errors to be expected from conventional mechanical scanners. In the scheme shown in the figure, optical switches are employed to alternately range to the target and reference legs of the synthetic Michelson interferometer. Since it is the difference between the measurements that is of interest, common mode phase-noise in the electronics and optics upstream from the beamsplitter (in particular, the scanning mirror), are cancelled. A means of removing, by calibration, path phase errors due to distortion in the optics is also possible. Target address biases due to imperfect optics can be determined through the use of a concave reference or calibration sphere centered on the output lens focus. Accuracy of the scanning system thus becomes dependent on the accuracy with which the calibration sphere is fabricated, and is not limited to the accuracy of the system optics.

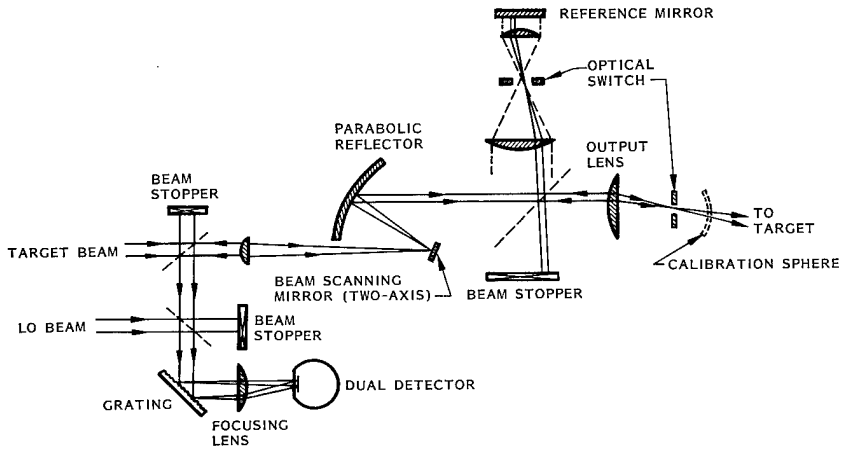


Figure 9. ADS beam pointing system

In outfitting an ADS system to meet the requirements of a specific mission, considerable emphasis will be placed on integrating the sensor into the structure to be sensed and controlled. This problem tends to be unique to a given undertaking, and hence is to be approached as a separate issue for each application.

In conclusion, it is fair to say that ADS is a new sensor system with a unique capability – a key sensor around which new, large-scale deployments for space as well as ground based missions can be configured. Experience indicates that where new sensors and future missions are related, today's capability tends to become tomorrows requirement.

Acknowledgement

The authors take pleasure in acknowledging the participation of individuals whose assistance made this work possible: studies and early concepts, J. S. Osmundson; analyses, L. Davis; interface hardware and computer technology, W. M. Wells and R. B. McIntosh; Bragg-cell development and optical subsystems, Dr. C. C. Huang; breadboard assistance, J. R. Hodor. Support for this work was provided by the Advanced Projects Agency of the Department of Defense under Subcontract 9540-A-0002 to ITEK Corporation. We are particularly indebted to Lt. Col. W. J. Cuneo, Jr. and T. G. Pitts for encouragement and guidance.

References

1. Tilford, C. R., Analytical Procedure for Determining Lengths from Fractional Fringes, Applied Optics, Vol. 16, No. 7, pp. 1857-1860, July 1977.
2. Bourdet, G. L. and Orszag, A. G., Absolute Distance Measurements by CO_2 Laser Multiwavelength Interferometer, Applied Optics, Vol. 18, No. 2, pp. 225-227, January 15, 1979.
3. Mocker, H. W., Rotational Level Competition in CO_2 Lasers, IEEE Journal of Quantum Electronics, Vol. 4, No. 11, pg. 769, November, 1968.
4. Schiffner, G., Prediction of CO_2 Laser Signatures, IEEE Journal of Quantum Electronics, Vol. 8, No. 12, pg. 880, December 1972.
5. Edlen, B., The Refractive Index of Air, Metrologia, Vol. 2, No. 2, pp. 71-80, February 1966.
6. Owens, J., Optical Refractive Index of Air, Applied Optics, Vol. 6, No. 1, pp. 51-59, January 1967.

Selected five color operation of a CO_2 laser

N. E. Buholz
Sensor Technology Department, Lockheed Missiles & Space Company, Inc.
O/62-12, B/151, P.O. Box 504, Sunnyvale, California 94086

Abstract

A simple CO_2 laser can be caused to exhibit a specific sequence of five rotational lines or "colors" with cavity length tuning. The sequence, useful in multicolor interferometric distance measurement, is predicted to appear only near certain discrete laser cavity lengths. Three different length breadboard lasers are described in which a desired sequence is found.

Introduction

In a recent letter[1] it was shown that a simple CO_2 laser can be caused to exhibit a desired sequence of three rotational-vibrational lines, or "colors" with cavity length tuning. Simple length adjustments produced the desired sequence as part of an appropriate signature. The two color ambiguity length[2] was used as a guide to determine the amount of length adjustment required. In this letter we show that a much more restrictive signature may be obtained in which five specific colors are caused to appear in a selected sequence. As shown previously[3,4] and here also, the specific color sequence is very useful in multicolor interferometry applications. As a consequence of this application, the laser length is restricted to a very few of the large number of $\lambda/2$ (λ optical wavelength) increments possible in laboratory sized (1 m) lasers. In the following, the connection between the application and laser length restriction is explained, a simple technique to predict the correct length is presented, and three breadboard lasers of different length are described in which specific color sequences have been obtained.

Color sequence selection

In general terms, the desired sequence of rotational lines (10.4 micron band) with cavity length tuning has the form

$$P(x)\ R(y)\ P\binom{x+2}{x-2}\ R\binom{y+2}{y-2}\ P\binom{x+4}{x-4} \qquad (1)$$

or the same with R and P interchanged. Here, x, y represent numbers like 18, 20, etc. The value of this sequence arises from three main characteristics. First, having R and P lines alternate means that simultaneous oscillation always occurs with an R line and a P line together. This allows convenient line separation for laser stabilization purposes (to be discussed in a future letter). Second, with simultaneous R/P oscillation, data from the distance measurement interferometer[4] can be taken simultaneously on both optical fringes, in effect reducing the R/P ambiguity length measurement sensitivity to vibration. A significant reduction in the laser frequency stabilization requirements is also achieved. Finally, the sequence (1) in which we have adjacent P lines, and adjacent R lines, allows the formation of several higher order "synthetic" wavelengths which extend the ambiguity length by orders of magnitude beyond that of the individual R/P pairs. To illustrate how this is accomplished, we reproduce here in Figure 1 a wavelength "pyramid" similar to Figure 7 of Reference 4. The base of the pyramid shows line center frequencies[5] of the sequence used in the system of Reference 4. The next line up has the difference frequencies and corresponding wavelengths of the four R/P pairs. Higher levels show the longer synthetic wavelengths obtained by taking combinations of the R/P pairs and combinations of the synthetic wavelengths. At the top we have the longest synthetic wavelength of approximately 40 cm. It is observed that any sequence of the form (1) produces a similar structure having approximately the same values. Thus, as indicated in Reference 3, the form of the synthetic wavelength hierarchy is primarily a characteristic of the CO_2 rotational energy structure, and only a correct cavity adjustment is required to enable a laser to output the desired wavelength sequence. It is recognized, of course, that operation in R/P pairs implies that the actual oscillation frequencies are off line center by some amount consistent with the cavity c/2L interval (c light velocity, L cavity length). This perturbation is reflected in slight adjustments of all the numbers up through the pyramid.

As the initial step in the adjustment of the laser length to produce a sequence such as (1), we note that all of the oscillation frequencies must be divisible an exact integral number by the Fabry-Perot interval c/2L, where a slightly different L is required for each oscillation frequency. Exact determination of each L value however, is neither necessary nor desirable. It is convenient rather to work with the second layer in the pyramid consisting of the R/P pair difference frequencies. We note that if line center differences are used, as shown, a fixed c/2L interval can be divided into all four R/P pairs, and some L values exist for which all divisions yield integers plus small fractional remainders. The remainders are close to zero or one and are indicative of the separation of the actual oscillation frequencies from the respective R/P line centers.[1] For example, if a remainder is 0.1, and we have an L = 1 m laser, the actual oscillation frequency interval differs by 15 MHz from the line center difference. The 15 MHz is split between the two R/P lines according to their gain/loss ratios. Thus, lengths where one should begin to look for a desired signature are those which result in integers plus small

remainders when c/2L is divided into the four R/P pair frequencies. Note that asking for such integral division is equivalent to setting L so that all four R/P ambiguity wavelengths "fit" a respective integral number of times into the cavity round trip length, 2L. These divisions can be made simply and rapidly using a desk top computer. It is convenient, however, to evaluate within a few millimeters where to begin such a calculation. For this information, we look to the higher order synthetic wavelengths in the pyramid. Since each level is composed of difference frequencies between members of the next lower level, each difference frequency in a given level must also be integrally divisible by c/2L. Equivalently, each corresponding synthetic wavelength must fit a whole number of times into the cavity round trip length 2L. This holds true for each level in the pyramid, and thus for the top, where we have the longest synthetic wavelength λ_{sl}. Therefore, values of laser cavity length at which to begin to look for the proper signature are given by

$$2L = n \lambda_{sl} \tag{2}$$

where n is an integer. Thus, the shortest laser in which the desired signature is to be found is approximately 20 cm long. In the long length limit, higher gain/loss lines tend to dominate because of the decreased mode spacing c/2L, and the signature would be unobtainable. The procedure for finding a desired signature may therefore be described as follows. First, determine a candidate laser length using Eq. (2) as well as specifications based upon total output power, beam size, etc. Next, form a division of the R/P line center difference frequencies by c/2L, incrementing L a small amount (100 microns, for example) for each set of four divisions. Look for quotients which have all four remainders near zero or one. Corresponding values of L are good candidates at which to start examination of signatures in the operating laser. The operating laser must be incrementally length adjustable with reasonable accuracy over a range of several millimeters. Finally, make incremental length adjustments, and observe the signature following each adjustment. Use the concepts described in Reference 1 to estimate the amount of adjustment necessary, and continue until the desired sequence is obtained.

Experimental results

Experimental work in support of these concepts has been limited to the construction and operation of three lasers having lengths corresponding to n = 5, 3 and 2 in Eq. (2). In each case, the procedure outlined above was used to obtain a sequence having the form of (1). No attempt has been made to find a similar sequence at intermediate lengths. A brief description of the lasers and sequences obtained follows (see Figure 2). The n = 5 laser has a nominal length of 1 meter. It is utilized in the system of Reference 4, and was also discussed in Reference 1. The tube has approximately 80 cm active length, with a 1 cm inner bore diameter. The two cavity mirrors have radii of R = 10 m and R = ∞ respectively. A TEM_{00} aperture with thin wire polarizer is included. Twenty watts TEM_{00} is nominally available at the flat output mirror. The laser utilizes pre-mixed flowing gas (the CO_2 mix) of 4.5:13.5:82 CO_2:N_2:He, with a pressure of 23 Torr and flow rate of 0.2 liters/min. It exhibits the following sequence, which fills most of a $\lambda/2$ cavity length sweep, in order of increasing frequency:

$$P(18)\ R(16)\ P(20)\ R(18)\ P(22) \tag{3}$$

The laser has been operated with this sequence in the system of Reference 4 over the past 5 months with no adjustment of either mirror tilt or cavity length. It was noted, however, that through an inadvertent change in gas mixture to one which was labeled 4:17:79 CO_2:N_2:He, a weakening of the R lines relative to the P lines occurred, and the sequence was nearly disrupted.

A sealed tube configuration was used in the n = 3 laser. This laser, 60 cm long, had mirrors with radii of R = 5 m and R = ∞, respectively. A TEM_{00} aperture was also used. The tube, with 30 cm active length and 1 cm inner bore diameter, produced 1.6 watts TEM_{00} at the output (flat) mirror. A PZT on the 5 m mirror provided 5 micron length adjustment. The following sequence was obtained in part of a 5 micron length sweep:

$$P(22)\ R(18)\ P(20)\ R(20)\ P(18)\ R(22) \tag{4}$$

Notice that two sequences of the type (1) are found here, one having three R lines, the other, three P lines. It is also observed that either sequence in (4) occurs over a smaller portion of the 5 micron length sweep than does the sequence (3). This is reasonable because of the increased mode spacing c/2L over that of the n = 5 laser.

The n = 2 laser, 40 cm long, utilizes a flowing gas tube like that of the n = 5 laser, but is shorter (active length 18 cm). It also has mirrors with radii of R = 5 m and R = ∞. A TEM_{00} aperture/thin wire polarizer is used here as well. Power output from the flat mirror is approximately 0.6 watts TEM_{00}. The sequence for this laser (5) occupies even less of the 5 micron signature than does either in (4) for the 60 cm laser.

$$R(18)\ P(18)\ R(16)\ P(20)\ R(14) \tag{5}$$

Summary

A method has been presented for adjusting the cavity length of a simple CO_2 laser so that it exhibits a desired sequence of rotational lines with cavity length tuning. The method, based upon multicolor interferometry concepts, involves an initial length estimate and lab-set final adjustments. Three breadboard lasers were described in which a desired sequence has been obtained.

References

1. N. E. Buholz, IEEE J. Quantum Electron. (to be published).
2. Charles R. Tilford, Applied Optics 16, No. 7, July 1977.
3. G. L. Bourdet and A. G. Orszag, Applied Optics, Vol. 18, No. 2, pp. 225-227, January 15, 1979.
4. C. W. Gillard, N. E. Buholz and D. W. Ridder, Proc. SPIE, Symposium '80 East #228-07.
5. Gerhard Schiffner, IEEE J. Quantum Electron., Vol QE-8, pp. 877-882, December 1972.

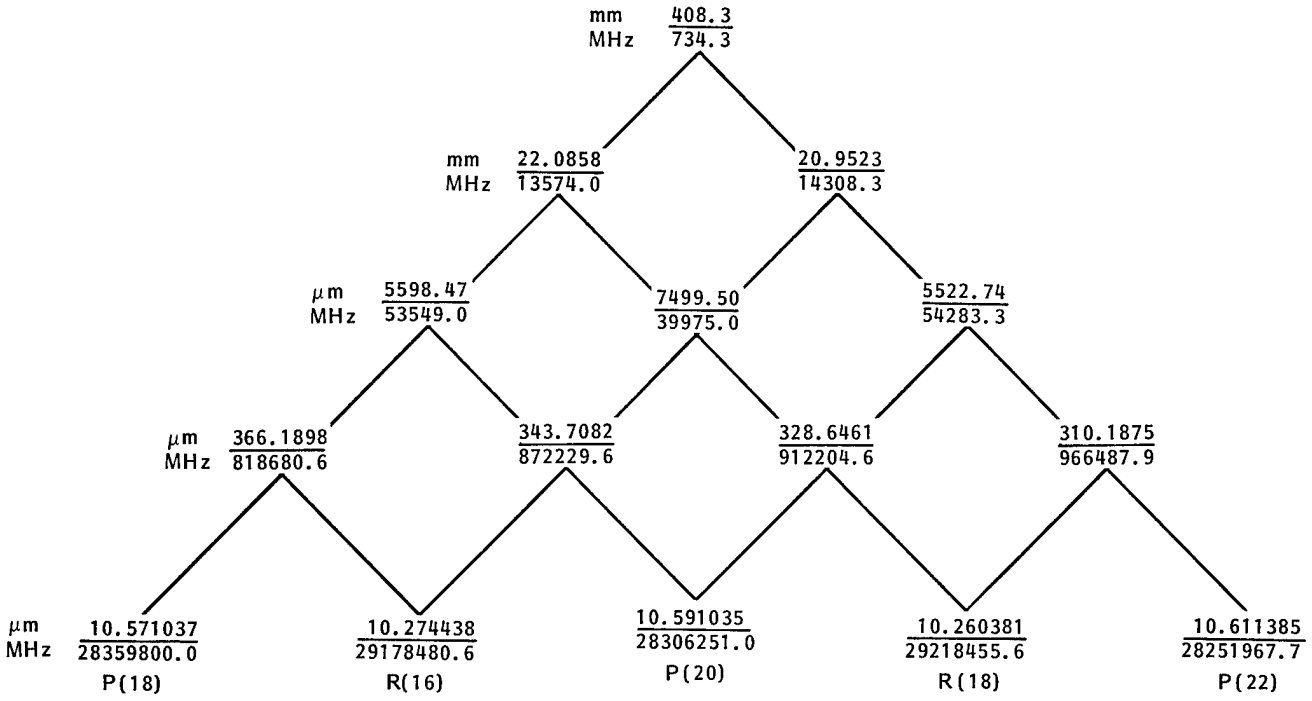


Figure 1. Synthetic wavelength pyramid.

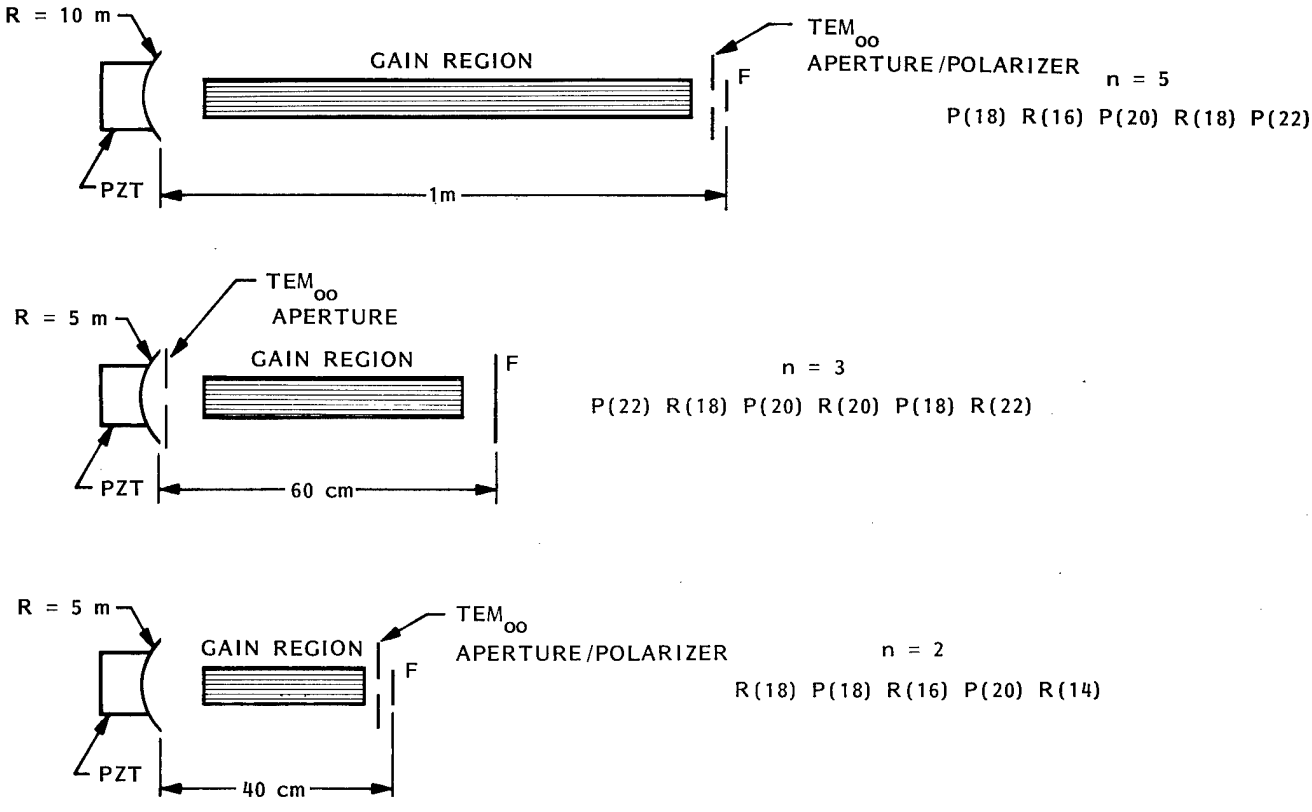


Figure 2. Breadboard lasers and line sequences.

A perspective on large space structure control

C. S. Greene, R. E. Pope
Honeywell Systems and Research Center, Minneapolis, Minnesota 55413

Abstract

In this paper a control engineer's point of view of the Large Space Structure (LSS) problem is presented. A definition of a "large" space structure from a control engineer's standpoint is presented along with a discussion of the various factors which drive the control problem. The interaction of the controller with the structure to be controlled is explored to point out the role of each disipline in solving any practical LSS problem. The ways in which they can contribute to a successful system are presented. Finally, the role of the optics engineer as both system user and technology provider is discussed.

Introduction

Future DOD, NASA, and commercially sponsored space missions will require spacecraft much larger than today's spacecraft. New high gain antenna systems with diameters on the order of thousands of meters are being proposed. A 26 mile Space Power System is also being actively considered by NASA. While size alone will present problems for future spacecraft control system designers, the most challenging problems will result when size is viewed in terms of weight. For space missions, weight means cost. As is done today, future spacecraft and payload designers will make every attempt to minimize weight. Additional constraints will be imposed by the requirement to use the Space Shuttle as the launch vehicle thus confining the spacecraft and payload to the size of the Shuttle payload bay.

The system specifications are at the same time, increasing. Antennas are getting larger with narrower beams thus requiring finer pointing. Higher operating frequencies are causing surface quality specifications to increase. The combination of large size, low weight, tighter specifications and shuttle payload characteristics will result in future spacecraft which, structurally, are much less rigid than todays spacecraft.

In the sections that follow we will discuss the impact of this reduced structural rigidity, or in other words, increased structural flexibility on the control system design problem. In Section II, we will elaborate on future trends with respect to the key control system design drivers. Section III will present some of the specifics of spacecraft control which would support an optical system mission. We will characterize the control problem in terms control designers understand and appreciate in Section IV. In Section V, we will demonstrate the realization of the control problem in two design example: an academic beam structure example and the Power Extension Package control problem. Finally, we will summarize in Section VI with our current perspective of large space structure control and the future technology impact.

Future trends in large space structures

Four trends in LSS have significant impact on the control problem. Structures are getting larger, more flexible, have greater accuracy requirements for pointing and surface quality and have larger disturbances.

Size - Figure 1 is a plot which demonstrates the driver forcing size to increase. The major point is that resolution and power increase for a given operating frequency as the diameter increases. Thus, larger antenna's permit operation at higher frequencies and/or higher gain/resolutions. The figure includes some current and proposed spacecraft.

Although Figure 1 is in terms of the drivers for radio frequency systems, the same drivers apply to optical ones as well. The drivers of increased power and resolution still force size to increase. A further discussion of the special features of optical systems is presented later.

Flexibility - Every pound sent into orbit costs money. Thus, there is a large incentive to keep weight to a minimum. Furthermore, when the space shuttle enters operation, spacecraft will no longer need to be as strong to stand the launch environment and can be sized only to withstand the relatively (from a structural loads standpoint) benign environment of

space. Thus, the weight penalty is even now forcing spacecraft to be as flexible as possible.

Accuracy - Science always pushes for greater accuracy from its instruments. Thus one can see a constant pressure for better pointing accuracy. This is especially true in the optics area. An instrument capable of fine resolution is useless if it cannot be pointed accurately.

Disturbances - In any control design, one of the significant drivers to the design is the disturbance environment. If the disturbances are small, little control energy is needed. Figure 2 shows some typical disturbances affecting a spacecraft. Note that in many cases, large disturbances are produced by the internal machinery of the spacecraft such as reaction wheels and control moment gyros (CMG). Not shown are other disturbances such as those caused by orbit maintenance operations which frequently are required while maintaining specified pointing, etc. As spacecraft grow, and become more complicated, more equipment is being included in the spacecraft.

There is, of course, an interaction between these drivers. For example, large flexible spacecraft are more susceptible to thermal disturbances, which are expected to be the major control driver for many spacecraft. Thus, it is important to understand each driver and each aspect of the above trends.

Optics requirements

For many spacecraft missions, it is the optics engineer which will determine the system specifications which must be met. These specifications directly drive the control problem and thus need to be understood by the controls engineer.

Two general classes of specifications can be seen. First there are those dealing with pointing accuracy. This is usually specified by two parameters. The first can best be thought of as a bias due to various alignment errors. It arises largely from the inability to mount components on a spacecraft arbitrarily accurately. Large size, of course, will complicate the alignment problem. The second parameter, usually refered to as pointing stability, refers to the dynamic errors involved in pointing. These errors result from the external and internal disturbances being transmitted to the device being pointed.

A second general class, becoming increasingly important from a control standpoint, specifies the surface quality of a reflector. Being new, standard approaches to its specification are still evolving. However, an oft quoted specification is that the desired surface must be maintained to one-twentieth to one-onehundredth of a wavelength. This of course is an imcomplete specification in that such variables as correlation length of the surface errors (both spatial and temporal) have large effects on reflector performance. A standard useful approach to surface quality specification remains to be developed although much useful research is currently in progress.

Characterization of the control problem

Feedback control has but one function: to cause a system to perform within specified requirements in spite of incomplete or inaccurate knowledge of the system. This ignorance of system characterisics comes mainly from uncertainties in the system itself and from random disturbances. This is clearly seen in optical systems. Only when manufacturing tolerances are greater than the imposed specifications or when disturbances are large does one consider the use of active control.

As we have seen, specifications on both the pointing accuracy of the overall system and on the quality of the optical/antenna surface have been increasing. In addition, the flexibility of the support structure has been increasing as can be seen in the reduction of the first bending frequency of the structure. These combined effects lead to an interaction between the structure and control system which, if ignored, can result in instability.

The effects of these can be seen in Figure 3. With conventional structures in which the control system bandwidth is significantly less than the first bending mode, control is a relatively straight forward task. The only requirement is that the controller be "rolled off" at high frequency (i.e., have loop gain less than unity) to avoid exciting the high frequency modes.

At the other extreme, when the control bandwidth overlaps with the bending modes of the structure, (see Figure 4) active structural control must be used. If a controller is designed which does not take into account the presence of the flexible modes, an unstable system will almost surely result. This overlap in frequencies is what most typically characterizes the LSS control problem.

Design examples

To illustrate these ideas, we will consider two examples. The first example, a simple one-dimensional free-free beam, has been used to "get smart on the basics". It has enabled us to understand the principals involved in LSS control without the complexity of more realistic structures. The second example is the Power Extension Package (PEP) solar power array for use with the space shuttle. This demonstrates the applicability of the basics to a real structure and also shows that the issue of LSS control problem is of current relevance.

Our initial studies have used a simple free-free beam as the structure to be controlled. The control goal was to augment the assumed natural damping of the beam (1%) to achieve 5% damping on the first mode. This was done with both a single sensor/actuator pair and with two pairs. The initial results have been published [1] and further work reported in internal memos [2,3] which will be published shortly. The major concusions will be summarized here.

1) Both the magnitude and frequency dependence of the inherent structural damping are critical. Light damping results in wide bandwidth controllers. Also, certain non-zero damping assumptions, commonly made in the control literature, result in infinite bandwidth requirements on the actuators, sensors and intervening computation. Damping is a very important driver to control system cost and yet is poorly understood and usually the most uncertain structural specification.

2) When certain sufficient conditions, given in [1], are met (as they almost always are in practice), multi-input, multi-output controllers can be simply designed which are insensitive to uncertainties in the modal characteristics of the structure. These results also specify the bandwidth over which sensors, actuators and computations need to operate and the frequency range of validity required for any verification model.

3) The insensitivity resulting from the controller designed in [1] is paid for, however because it has high bandwidth relative to current satellite controllers. If additional structural information (i.e. modal frequencies, damping ratios and influence coefficients for low frequency modes) is available, some reduction in bandwidth is possible at the expense of sensitivity to uncertainties in these parameters. Determining these parameters either pre-launch through modelling, scale model testing, etc. or post-launch with on-orbit testing and either adaptive control or reprogrammable control is currently an active reseach topic.

These control results have also been applied to NASA's Power Extension Package (PEP). PEP, shown in Figure 5, is a 13 foot by 270 foot solar array which can be carried compactly in the orbiter cargo bay, deployed on orbit to conserve expendables for the fuel cells and then restored in the payload bay for reentry. While on-orbit, it is deployed and held in place by the orbiter's Remote Manipulator System. In operation it includes a sun sensor and two gimbals to automatically track the sun. The control task is to design the controller for these gimbals.

The frequency response for a model of PEP for the alpha gimbal is shown in Figure 6. In this model, the first bendng mode which has significant infuence on the transfer function (i.e., there are lower frequency modes which are only weakly observable) occurs around .08 Hz. For the disturbance environment specified, the minimum bandwidth is roughly .02 Hz with something closer to .1 Hz being preferred. Since it is not possible to roll the control loop off fast enough to keep the .08 Hz resonance peak below unity gain, the technology of LSS control must be used. This can be done by rolling the loop off slow enough to prevent instabilities. As predicted by the previous analysis, this required high bandwidth gimbal drive compared to previous space applications.

The extreme importance of damping is again seen. Controllers which are robust to modal data have been designed and the effects of damping on these designs assessed. Table 1 summarizes these effects by showing the frequency range over which all feedback components are required to be accurate. Note that for very small inherent damping ratios, extremely wide bandwidths are required.

Summary and conclusion

We have attempted to present a control engineers view of LSS. These large, highly flexible spacecraft impose some new requirements for control system design, principally because of the interaction between the control system and the structural frequencies caused by an overlap of control bandwidth and structural dynamics. This interaction, which could previously be ignored, is now a major design driver and is especially worrysome in systems employing optical elements since they require very fine pointing, and surface control.

We have explored the implications on control design through two design examples, a free-free beam and a large solar array. These have shown inherent structural damping to be a critical parameter to control design and that, given sufficient damping, multi-input/output controllers can be designed which meet system specifications. A major results is that in general wide-bandwidth components are required to assure system stability compared to bandwidths of current satellite controllers.

Spacecraft control design is changing due to more demanding requirements. As a result, many of the standard notions about the role of the control system designer need to be re-examined.

References

1. Greene, C.S. and G. Stein, "Inherent Damping, Solvability Conditions and Solutions for Structural Vibration Control", 1979 Conf. on Decision and Control, Ft. Lauderdale, FL., Dec. 1979.

2. Greene, C.S. and G. Stein, "Inherent Damping, Solvability Conditions and Solutions for Structural Vibration Control", Honeywell Systems and Research Center MR12554, Minn., MN, April 1979.

3. Greene, C.S. and G. Stein, "Control of Large Space Structures: An Update", Honeywell Systems and Research Center MP12576, Minn., MN, Jan. 1980.

Damping (%)	Required Control Bandwidth (Hz)
5	.3
1	30
.5	10
.25	30
.1	120

Table 1 The damping bandwidth tradeoff for PEP

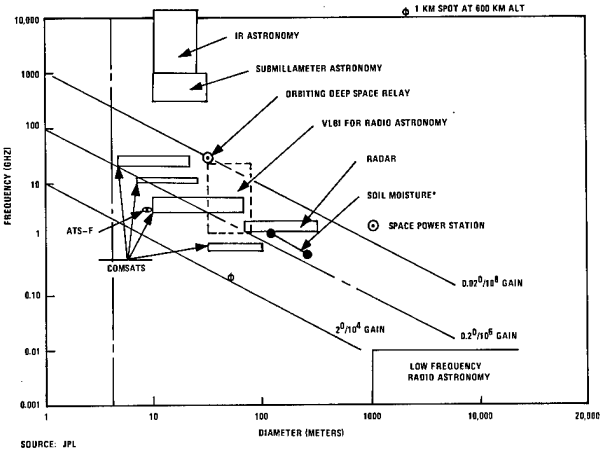


Figure 1. The size drivers for large antennas

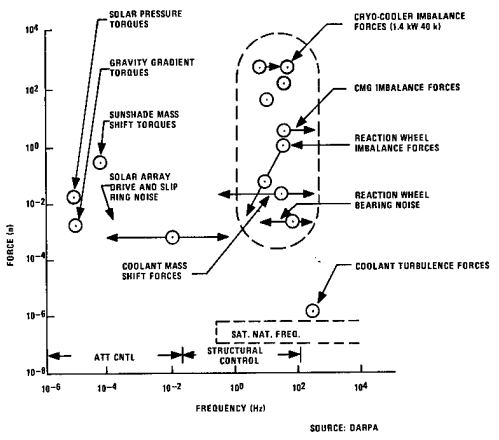


Figure 2. Typical disturbances

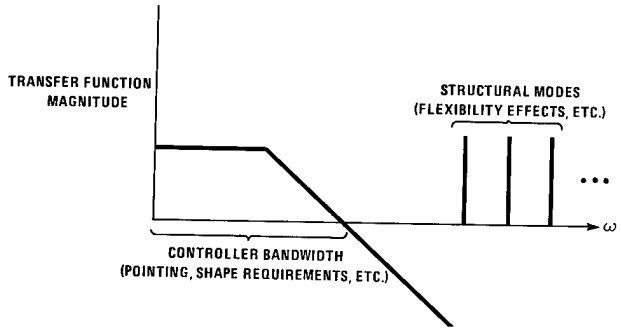


Figure 3. Conventional control

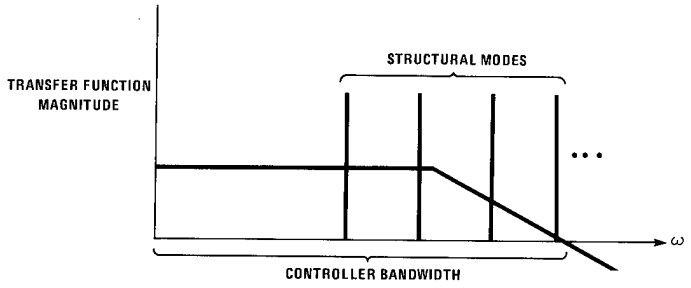


Figure 4. LSS flexibility control

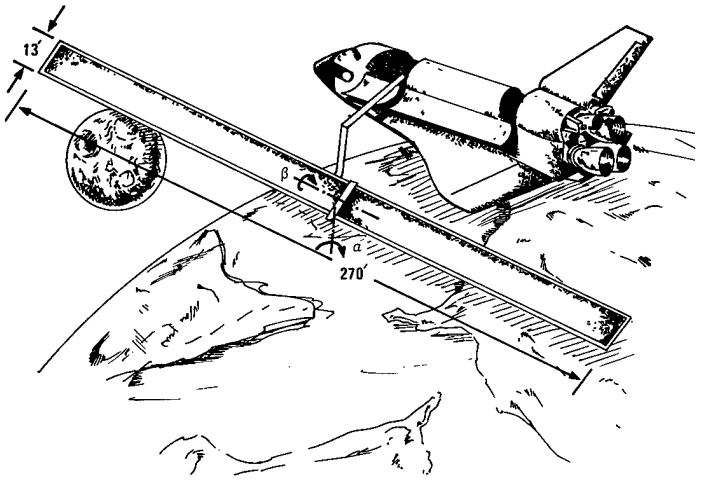


Figure 5. The Power Extension Package

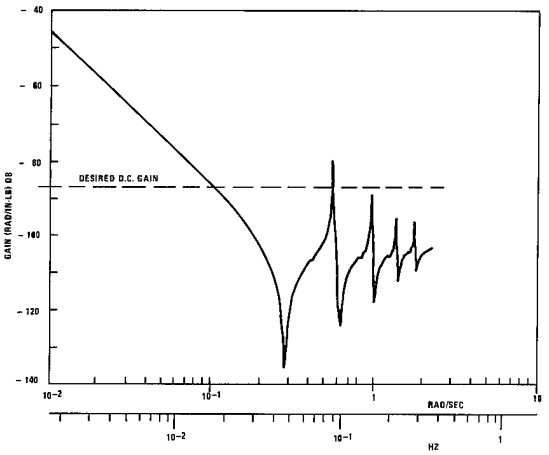


Figure 6. Frequency response for PEP alpha gimbal

ACTIVE OPTICAL DEVICES AND APPLICATIONS

Volume 228

SESSION 3

POTENTIAL NASA APPLICATIONS OF LARGE ADAPTIVE OPTICS

Session Chairman
Charles O. Jones
NASA Marshall Space Flight Center

Space astronomy to the year 2000: a preview of the possibilities

Jeffrey D. Rosendhal
Advanced Programs and Technology, Astrophysics Division
National Aeronautics and Space Administration, Washington, D.C.

Abstract

Following an overview of the role of space observations in contemporary astrophysics, the next generation of space-based observatories (the Space Telescope, the Gamma Ray Observatory and the Advanced X-Ray Astrophysics Facility) is described. Possible new directions which may be pursued in the 1990s are also discussed. These include the development of large flux collectors for use in astronomy in the ultraviolet, optical, infrared and millimeter wave portions of the spectrum and the development of space-based interferometers to carry out a variety of astrophysically important measurements. Many of these longer term programs will require substantial advances in optics, structures, and control technology.

Introduction and overview

As a result of the United States' program in space astronomy, during the past 25 years a revolution has occurred in Astrophysics. Rocket and satellite observations at ultraviolet wavelengths have shown that many types of stars are ejecting significant amounts of material at high velocities. Our concept of the interstellar medium—the gas and dust that pervades the vast distances between stars within our galaxy—has been totally altered. Prior to the space age, this medium was thought to be composed of two basic components: one dense and cold, and the other relatively warm. During the past decade, we have learned that this model is far too simple (Figure 1). Many different components of the interstellar medium have been discovered by our space instruments. We have seen very hot gas which is probably the result of tremendous collisions between the "normal" medium and shock waves from exploding stars. Moreover, we have learned that this medium is extremely clumpy and not the relatively uniformly distributed collection of gas and dust described in all the astronomy textbooks ten years ago. It is now known from both ultraviolet and X-ray observations that many types of stars possess tenuous high temperature outer layers and exhibit solar like activity but that the detailed behavior of these stars is very different from that expected on the basis of our current understanding of the solar chromosphere and corona. Discoveries of celestial X-ray and gamma ray sources have revealed the existence of entirely new types of celestial objects and have shown that we live in a Universe characterized by the routine occurrence of vast explosive events of unimaginable violence. Evidence has accumulated from X-ray measurements that suggests that a significant fraction of the matter in the Universe may exist in the form of very high temperature gas located between the galaxies. X-rays have also been observed to be emitted from objects ranging from nearby cool stars to the most distant quasars (Figure 2) and recent observations have shown that some of the most interesting types of astronomical bodies are emitting the bulk of their energy in the form of gamma rays (Figure 3).

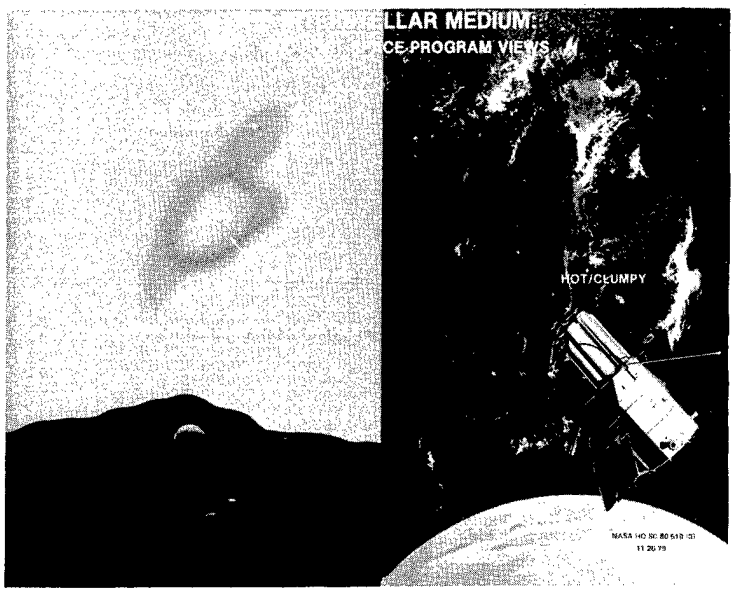


Figure 1. Changes in our understanding of the interstellar medium which have resulted from the space astronomy program.

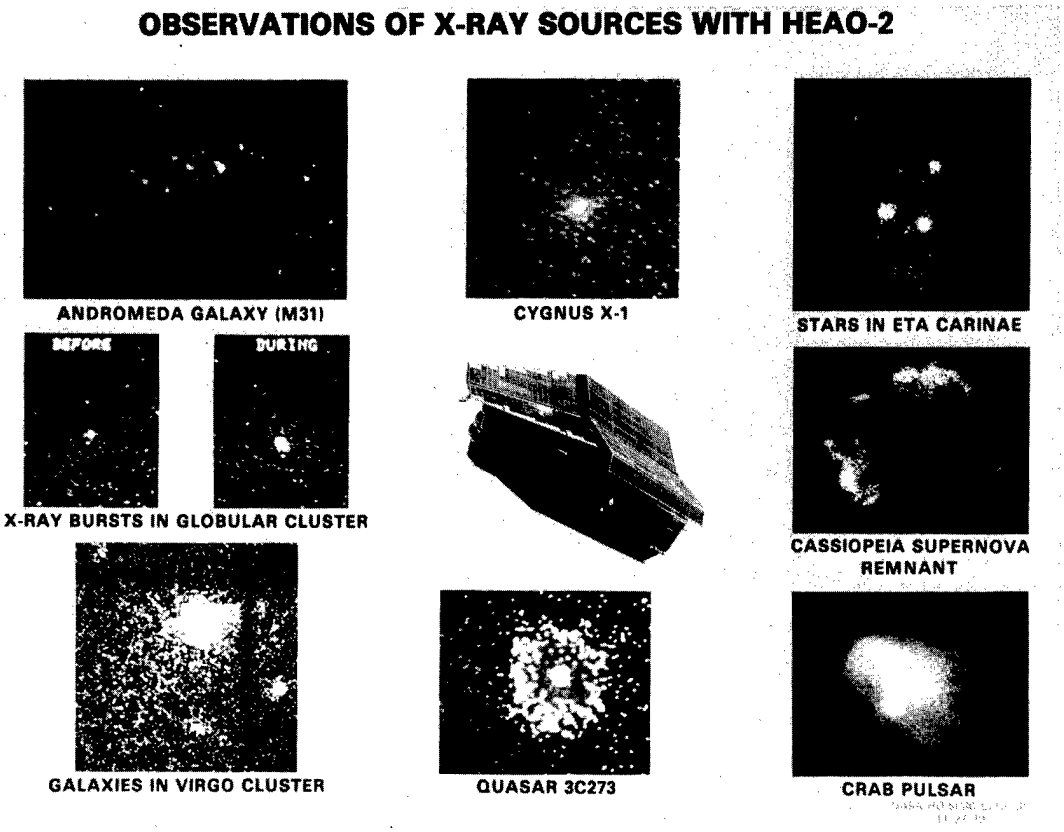


Figure 2. A collage of images of X-ray sources obtained with the HEAO-2 X-Ray Telescope. X-Rays have now been observed from objects ranging from nearby cool stars to the most distant quasars.

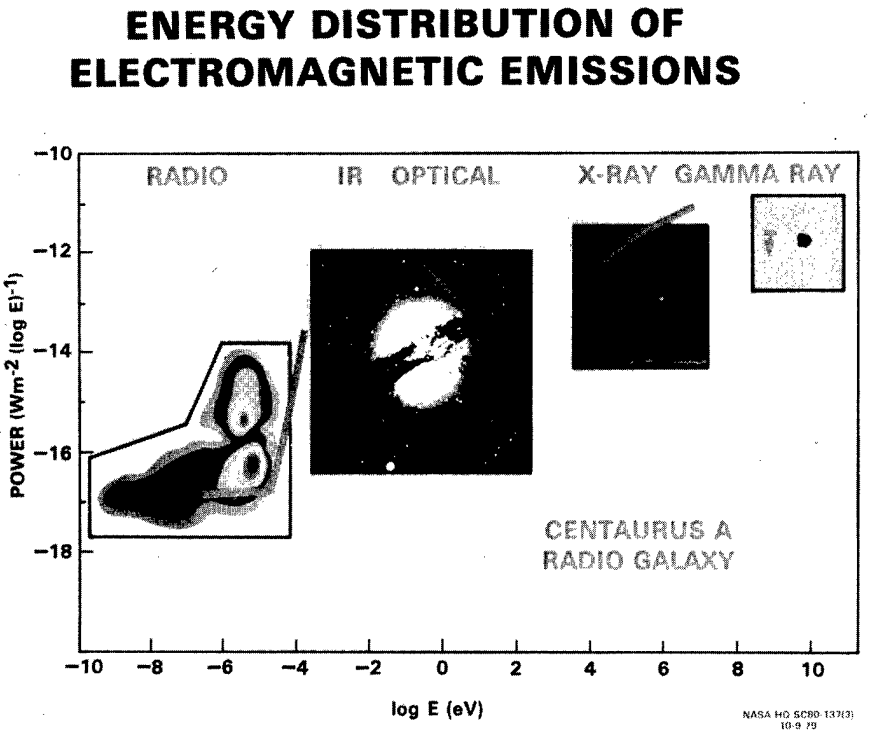


Figure 3. Distribution of energy emitted by the radio galaxy Centaurus A. The heavy lines show the observations while in the background is shown the appearance of the source in different parts of the spectrum. The gamma-ray appearance is as yet unknown but is shown as a dot because the gamma radiation probably originates from the nucleus of the source.

This revolution has taken place because we now have access to all the forms of electromagnetic energy emitted by celestial objects. For most of history, only visible light could be studied. Natural physical processes, however, give rise to characteristic radiations in a very much wider range of wavelengths (Figure 4). Each part of the spectrum may be thought of as a unique window to the Universe carrying unique information about physical conditions and physical processes. Infrared radiation reveals the presence of thermal emission from cool objects while ultraviolet or extreme ultraviolet radiation may result from thermal radiation from hot objects. Various types of violent events may lead to the production of X-rays and gamma rays. Since most types of radiation are blocked by the Earth's atmosphere, it was only when we had access to space and had developed the sophisticated instrumentation to take advantage of that access that we had a way of making many of the key observations.

The future astrophysics program is being designed to continue this on-going revolution in our understanding of the Universe. In this paper, we will briefly describe some of the potential major programs in space astrophysics which may be undertaken between now and the year 2000. Many of these programs will continue to exploit access to the full range of the electromagnetic spectrum. Others will exploit the fact that the turbulent atmosphere of the Earth not only blocks radiation but also distorts it. Use will be made of space to carry out observations at unprecedented angular resolution even at wavelengths accessible from the ground. Still others will exploit the space environment to measure primary cosmic ray particles or to perform unique experiments in fundamental physics which cannot be done in earth-bound laboratories. Some of these programs, particularly those likely to be started during the next three-five years, are at a relatively advanced stage of definition. For the longer term program, which is just now being defined, new technology developments will be required in many areas. Active optical systems and advanced control technology will be particularly important areas for development. It is obviously impossible to give an exhaustive description of every conceivable individual project. Rather, the objective here is to give an impression of possible trends and to indicate how the various pieces of the program fit together. More detailed technical and scientific discussions of some of the longer range possibilities are presented in other papers in this Symposium.

The near-term program (1981-1986)

The overall strategy for the Astrophysics program and the major initiatives which may be undertaken during the next five-seven years (as well as current and previous programs) are summarized in Figure 5. In this matrix, the first column lists a number of scientific disciplines or wavelength ranges (typically distinguished by differing instrumental requirements), the second column a representative scientific problem which can be studied at that wavelength range, and the remaining columns attempt to indicate the present degree of development in each area. The developmental stages which have been identified in the astrophysics program strategy are: (1) Preliminary surveys (low cost first cut) to see if anything is there and to detect gross features; (2) Initial all-sky surveys to acquire initial statistical information on source characteristics, determine approximate source locations, etc.; (3) Detailed study of individual sources or high sensitivity surveys; (4) Full-scale observatories; (5) Specialized follow-on and observatory support missions. In addition, individual missions may be developed to study specialized or unique problems requiring use of the space environment.

As is evident, different disciplines within astrophysics are at very different stages of development. At one extreme, the first extreme ultraviolet (100-900Å) sources have only recently been detected and the next step will be to carry out the initial all-sky survey. At the other extreme, building upon a long history of ground-based astronomy and upon a series of previous ultraviolet space astronomy missions, the

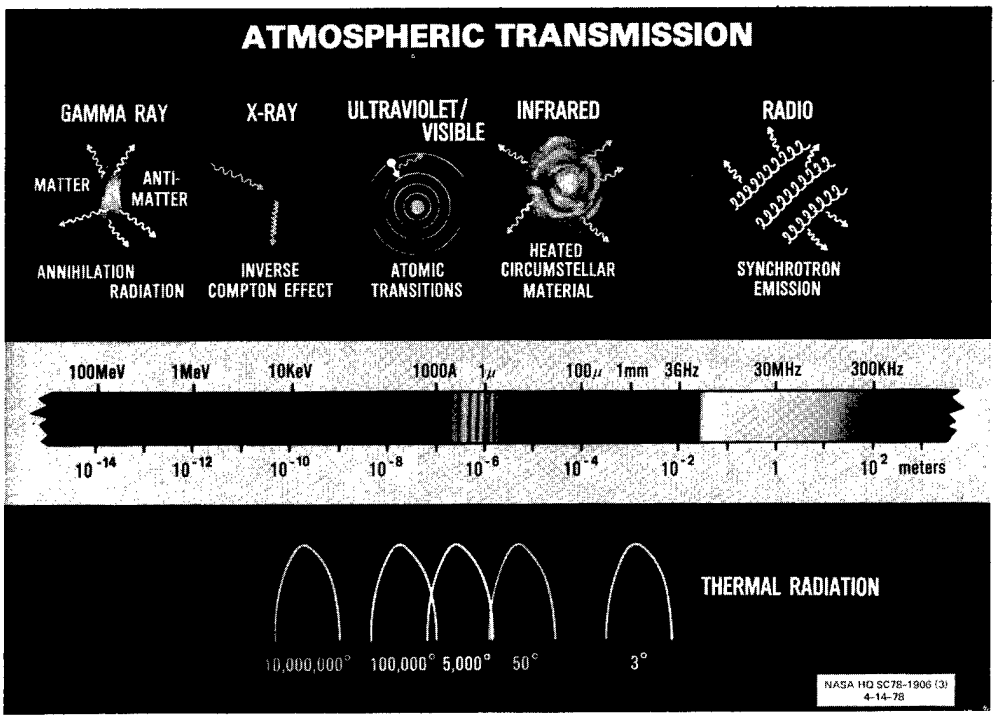


Figure 4. Obscuration of radiation emitted by astrophysical objects by the Earth's atmosphere. Objects at different temperatures or emitting radiation as the result of different physical processes emit radiation in different regions of the spectrum.

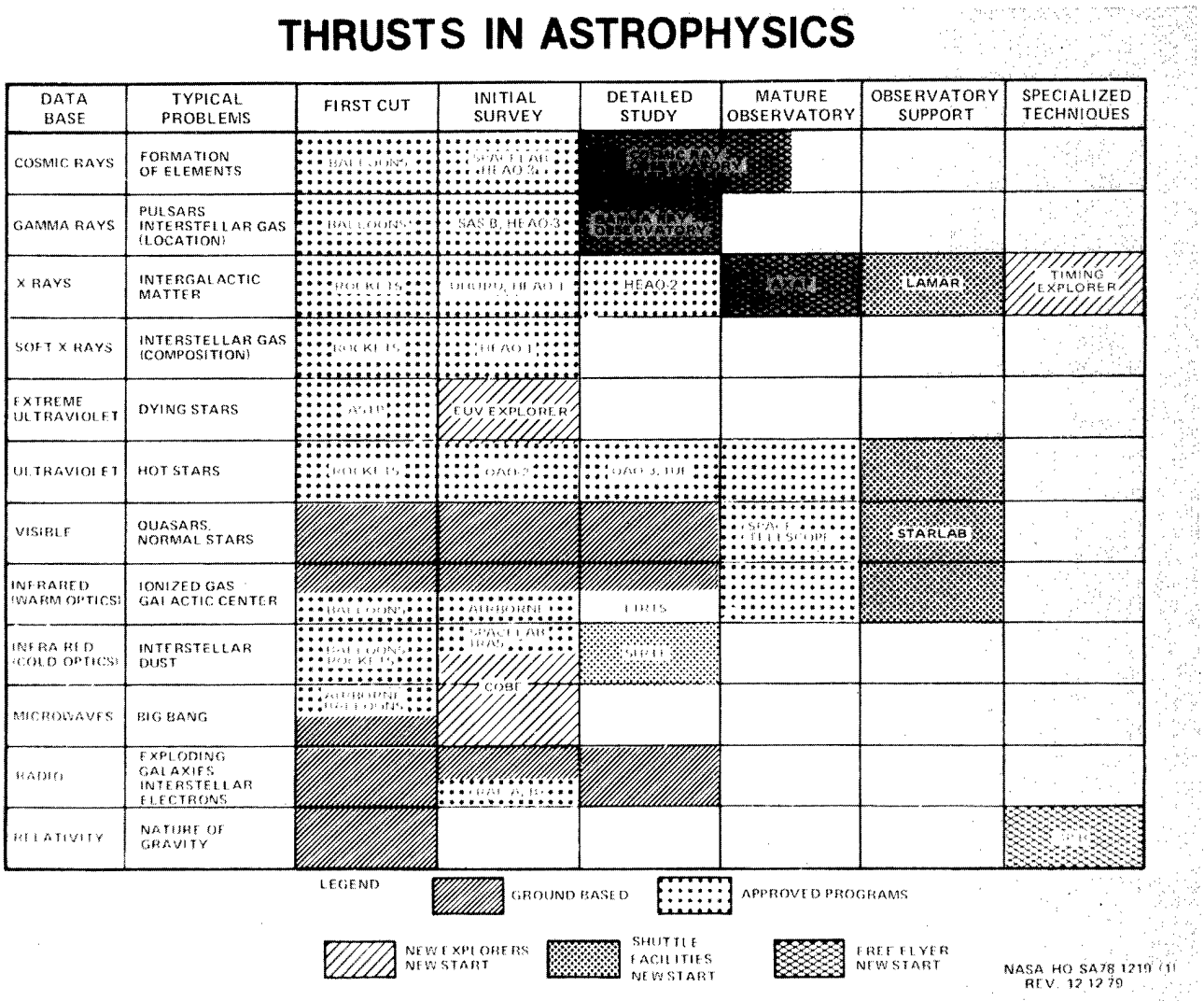


Figure 5. Summary of the stage of development of various subdisciplines in astrophysics. The matrix attempts to lay out how various past and future missions fit into an overall program strategy.

Space Telescope will move UV/Optical astronomy into the mature observatory stage. The discipline of X-ray astronomy is now ready to move into this stage as well.

The focus in this paper will be on the major initiatives. Other elements in the total program include smaller, specialized Explorer-class satellites and Spacelab instruments. Of the major programs shown in Figure 5, only the Space Telescope is now under development. The remainder of the projects are in the planning stage, with the Gamma Ray Observatory now being considered by Congress for a program start in Fiscal Year 1981.

The Space Telescope (Figure 6), to be launched by the Space Shuttle in December 1983, will constitute the heart of the United States Space Astronomy program until the end of the century. Its ability to cover a wide range of wavelengths, to provide fine angular resolution and to detect faint sources will make it the most powerful astronomical telescope ever built. A comparison of the theoretical expected performance of the Space Telescope with the performance of several satellite and ground-based telescopes is shown in Figure 7. It will be used to attack a wide variety of frontier problems in astrophysics, particularly in the areas of extragalactic astronomy and observational cosmology. Unlike previous astronomical satellites, which have had relatively limited lifetimes and instrumentation determined by the technology available many years prior to launch, the Space Telescope will be a long-lived observatory in space, accommodating a variety of scientific instruments which may be changed and updated (using the Space Shuttle) as both scientific priorities and instrument or detector capabilities evolve.

The Gamma Ray Observatory will move gamma-ray astronomy from the initial survey stage to the detailed study stage of its evolution. Gamma ray measurements reveal the explosive high energy nuclear and elementary particle processes that occur in the Universe (Figure 8). High sensitivity observations of gamma ray line emissions in supernovae and their remnants can provide direct evidence for the occurrence of nuclear reactions leading to element synthesis. Gamma rays produced by interactions of cosmic rays with the interstellar medium provide

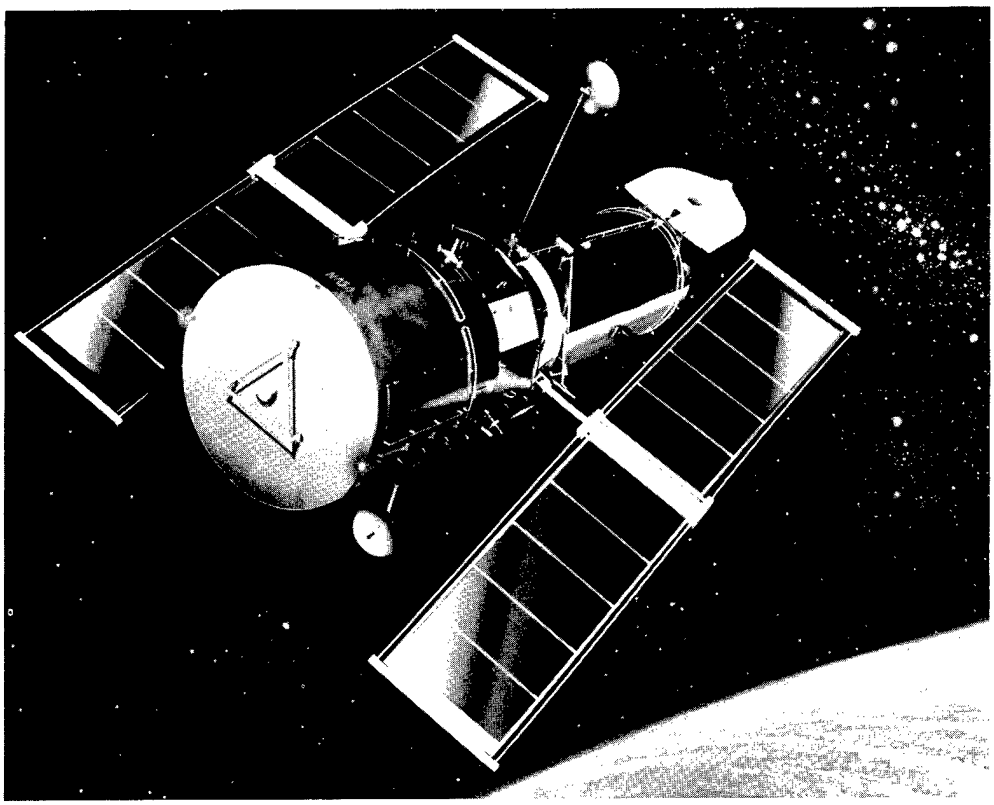


Figure 6. An artist's concept of the Space Telescope.

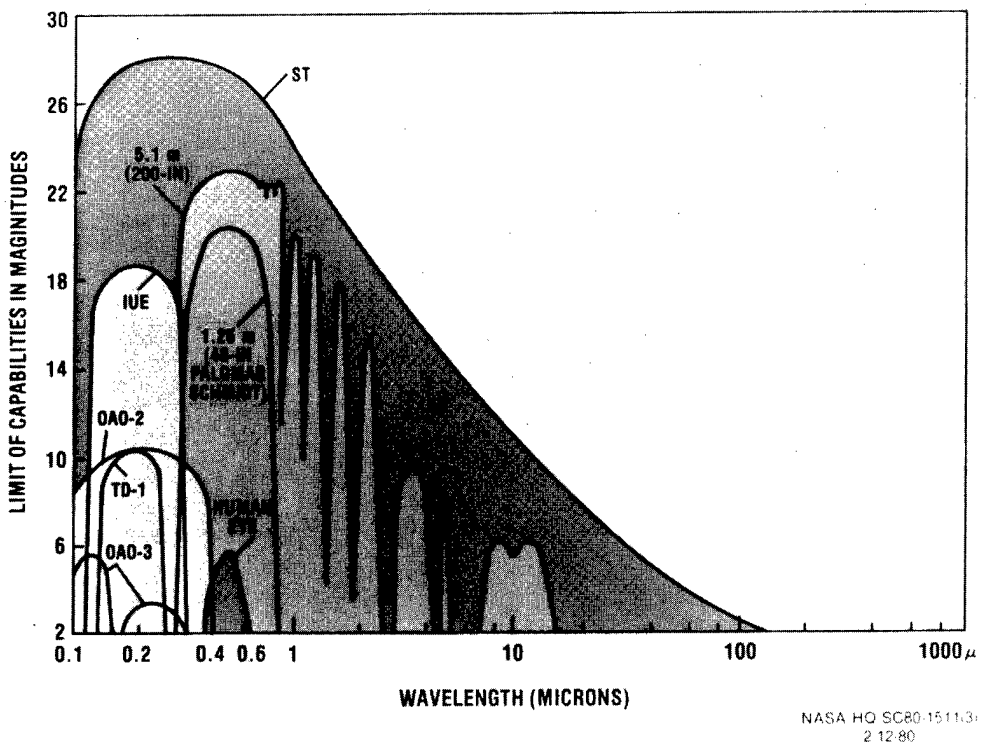


Figure 7. A comparison of the theoretical expected performance of the Space Telescope with the performance of several satellite and ground-based telescopes. The first instrument complement does not contain an instrument sensitive to wavelengths larger than 1.1 microns. However, it is expected that the Space Telescope ultimately will be used for infrared astronomy.

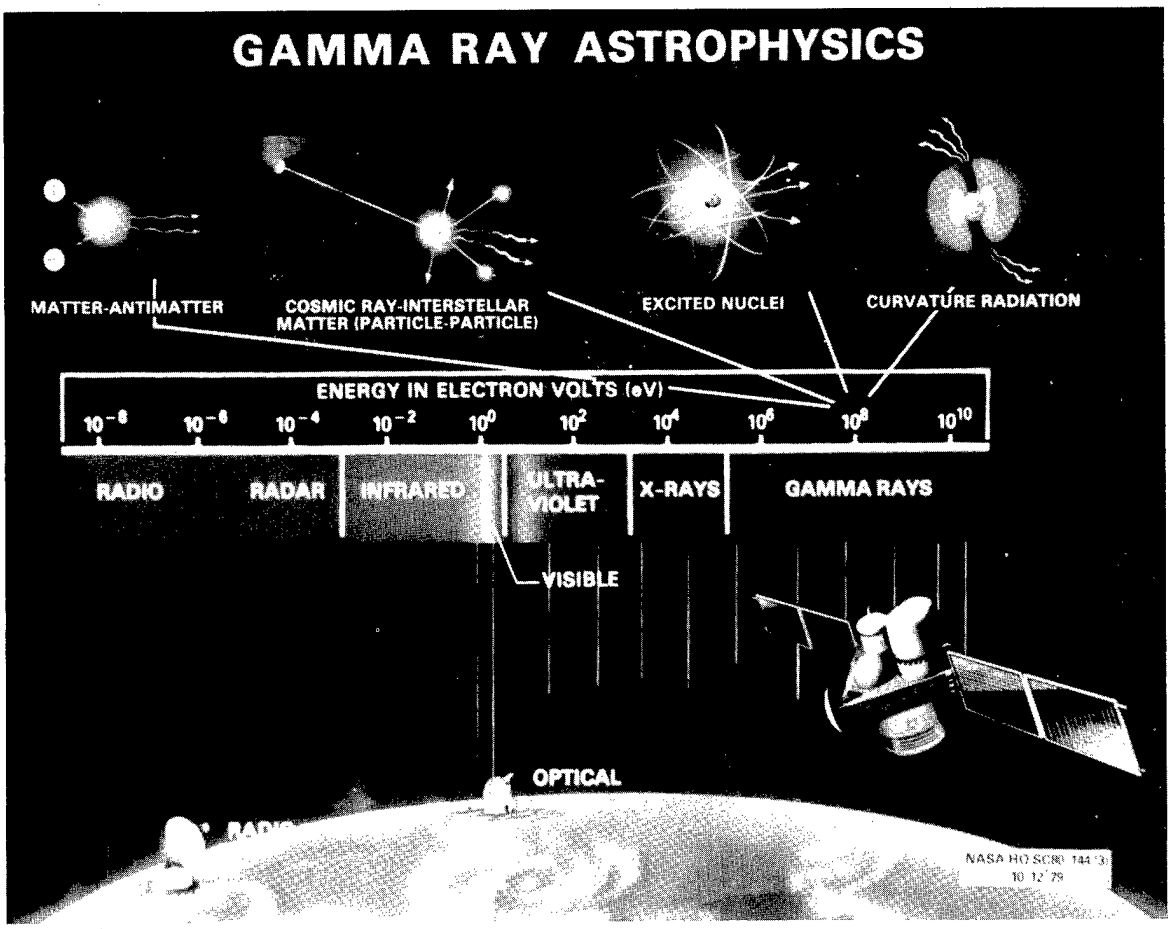


Figure 8. Unique physical processes which lead to the production of gamma rays in astrophysical sites. In many cases, gamma rays are the only way to observe such processes.

direct information on both the interstellar gas and cosmic rays. Gamma rays are an important component of the radiation emitted by many types of astrophysically interesting objects such as pulsars and active galaxies (See Figure 3). Because of their extreme penetrating power, gamma rays retain the detailed directional and temporal features imprinted on them at their birth, even if that birth was deep in regions opaque to visible light and X-rays, or was early in the history of the Universe. The Gamma Ray Observatory will be the first mission to study the full range of gamma ray phenomena covering six decades of the spectrum. It will observe sources at least ten times fainter than observed on previous gamma-ray missions. It will improve our knowledge of gamma ray source positions in the sky by a factor of 50-100 allowing searches for optical, X-ray or radio counterparts. It will explore previously unobserved portions of the gamma ray spectrum and it will be able to study the entire gamma ray sky. The scientific instruments to be developed for the Gamma Ray Observatory, the energy coverage of each instrument, and the energy ranges important for several astrophysical studies are illustrated in Figure 9.

Gravity Probe-B (Figure 10) will carry out a fundamental test of the theory of general relativity (Figure 11) by measuring the geodetic precession due to the motion of a gyroscope through a gravitational field (relativistic spin-orbit coupling) and the motional precession produced by the twisting of space due to the rotation of the Earth itself (relativistic spin-spin coupling). Detection of the latter effect would be an entirely new test of the theory. High precision measurement of the magnitude of these two effects (which according to general relativity are 6.9 arc seconds/year and .050 arc seconds/year, respectively, for a satellite in a 300 nautical mile orbit) should constitute an important step in discriminating between viable, competing theories of gravitation. Measurements of the necessary precision can only be made in the low gravity environment of space over a prolonged period of time. Work on the development of the technology for this experiment has been underway for more than 15 years during which time some extraordinary technical advances have been made particularly by the group at Stanford University. It now appears that the technology has progressed to the point where the mission is a serious candidate for flight.

The Advanced X-Ray Astrophysics Facility (AXAF-Figure 12) will take X-ray astronomy into the mature observatory stage of its evolution. The ubiquity of celestial X-ray emission has been well established by the HEAO missions, particularly by the HEAO-2 X-ray imaging telescope. The AXAF will be the direct descendent of the HEAO-2 mission. The focal plane of the grazing incidence X-ray telescope, which will be about 1.2 meters in diameter, will be capable of accommodating a variety of instruments to provide spectral and high spatial resolution data on quasars, galaxies, clusters of galaxies, and the intergalactic medium. AXAF will be launched, serviced and retrieved by the Space Shuttle and is also intended to be a long-lived National Facility. With four times the spatial resolution and at least twenty times the sensitivity of HEAO-2, AXAF will represent as large an advance in X-ray astronomy as the Space Telescope will be in optical astronomy. The increase in sensitivity is vital for making the next generation of detailed measurements on the X-ray objects discovered by the HEAO missions.

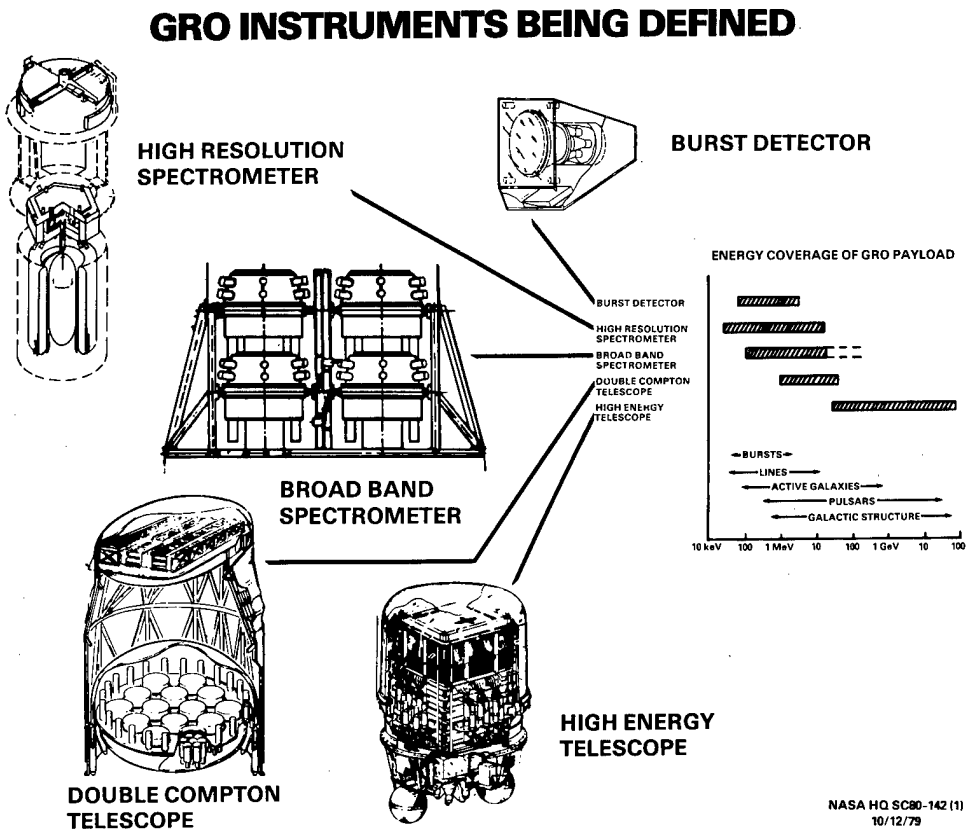


Figure 9. The GRO instrument complement.

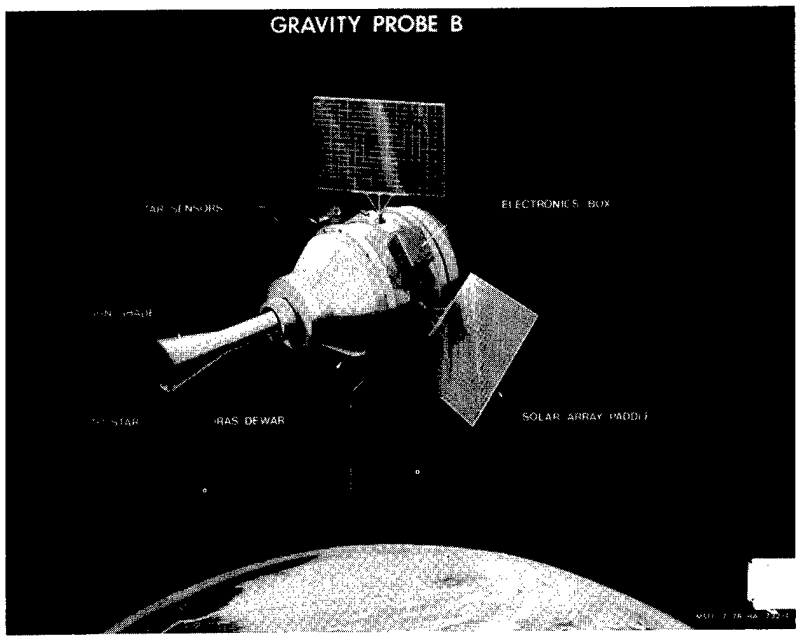


Figure 10. One possible concept of the spacecraft configuration for the orbiting gyroscope experiment Gravity Probe B.

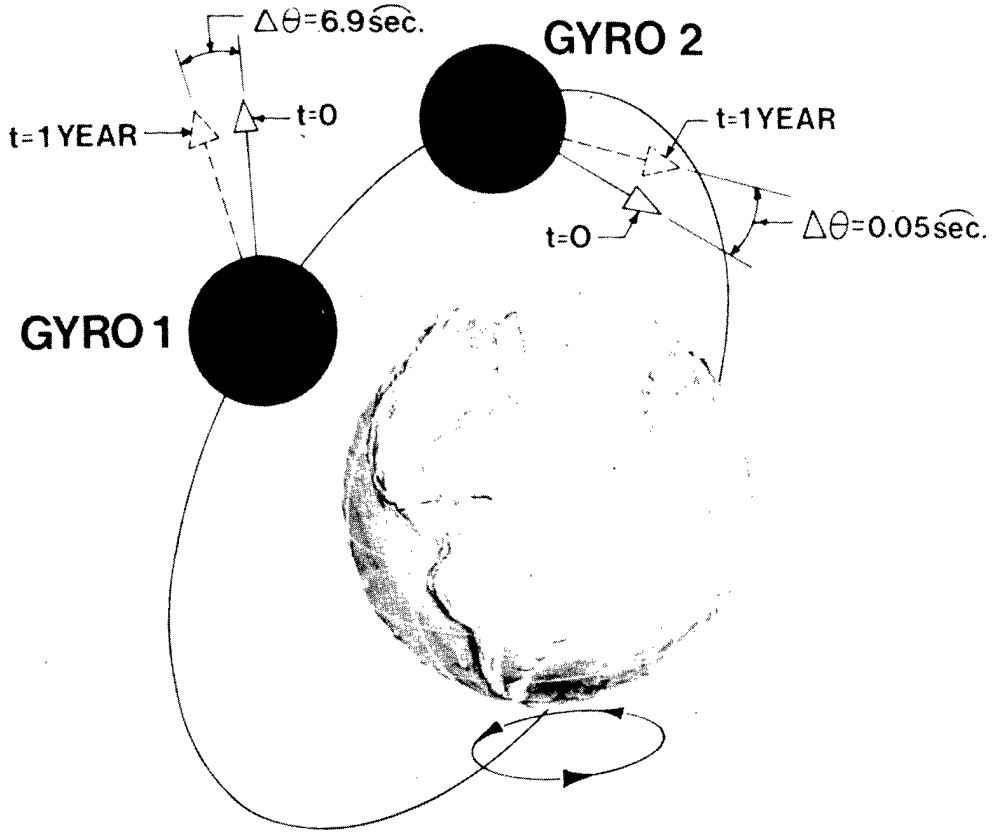


Figure 11. Relativistic motions of an orbiting gyroscope. In a polar orbit the two effects are at right angles to each other and would be most easily separated by properly oriented gyroscopes.

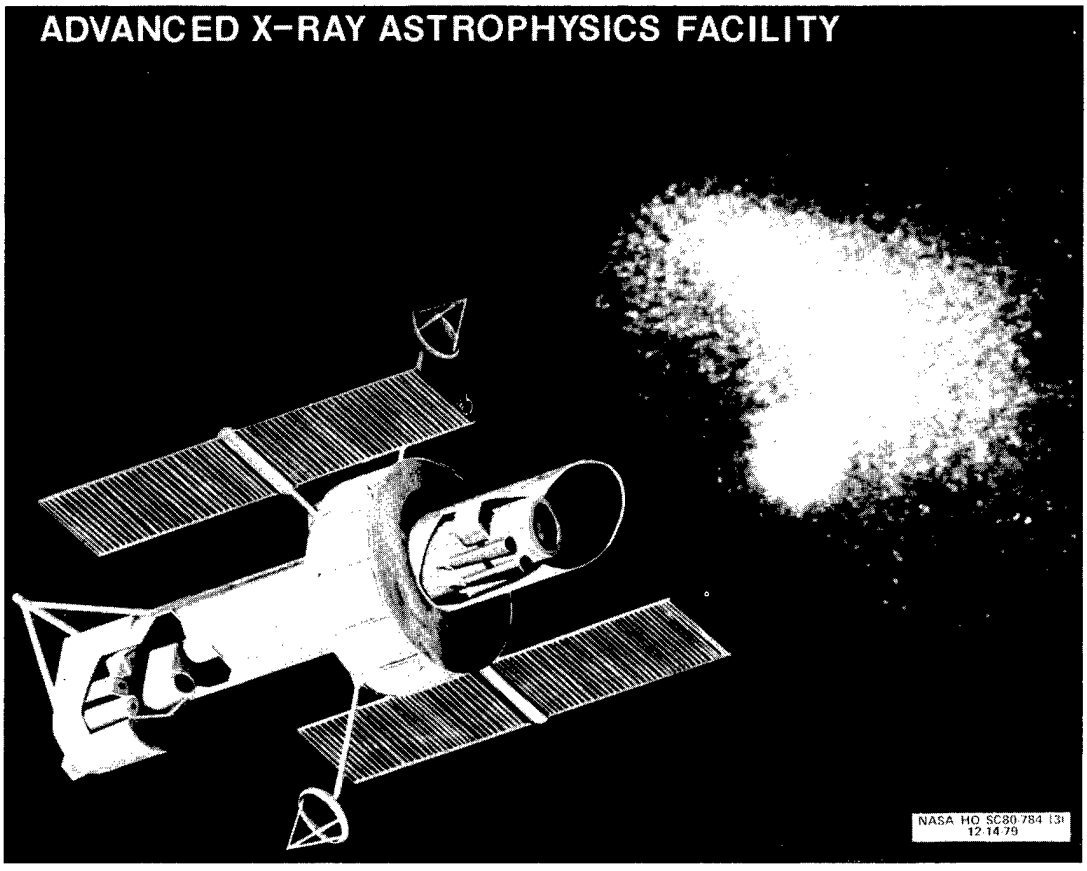


Figure 12. An artist's concept of the Advanced X-Ray Astrophysics Facility (AXAF).

The Cosmic Ray Observatory (Figure 13) will use large, massive instruments (which may be initially tested and used on Spacelab flights) to carry out measurements of cosmic rays of very high energy and very low flux. Such an observatory will provide the large area and long exposure time required for frontier investigations on the charge composition, energy spectra, arrival direction, and isotopic composition of cosmic ray nuclei. Other problems such as the existence of exotic particles (e.g.. superheavy nuclei, magnetic monopoles and antinuclei) may also be studied.

A possible (and probably optimistic) scenario for flight of all the elements in the current astrophysics 5-Year Plan is shown in Figure 14. As is evident, many of these programs will not be launched until the late 1980s, and some of the major facilities will be in use well into the 1990s. Programs which may be started in the second half of this decade or later will take us rapidly toward the year 2000.

A longer term outlook (1986-2000)

Because the process of planning and defining new missions and developing the technology to implement these missions frequently requires a decade or more, it is necessary to continuously think in terms of a comprehensive long-term program. We have identified and will continue to identify additional scientific requirements and missions needed to carry out a broad program in Space Astrophysics throughout

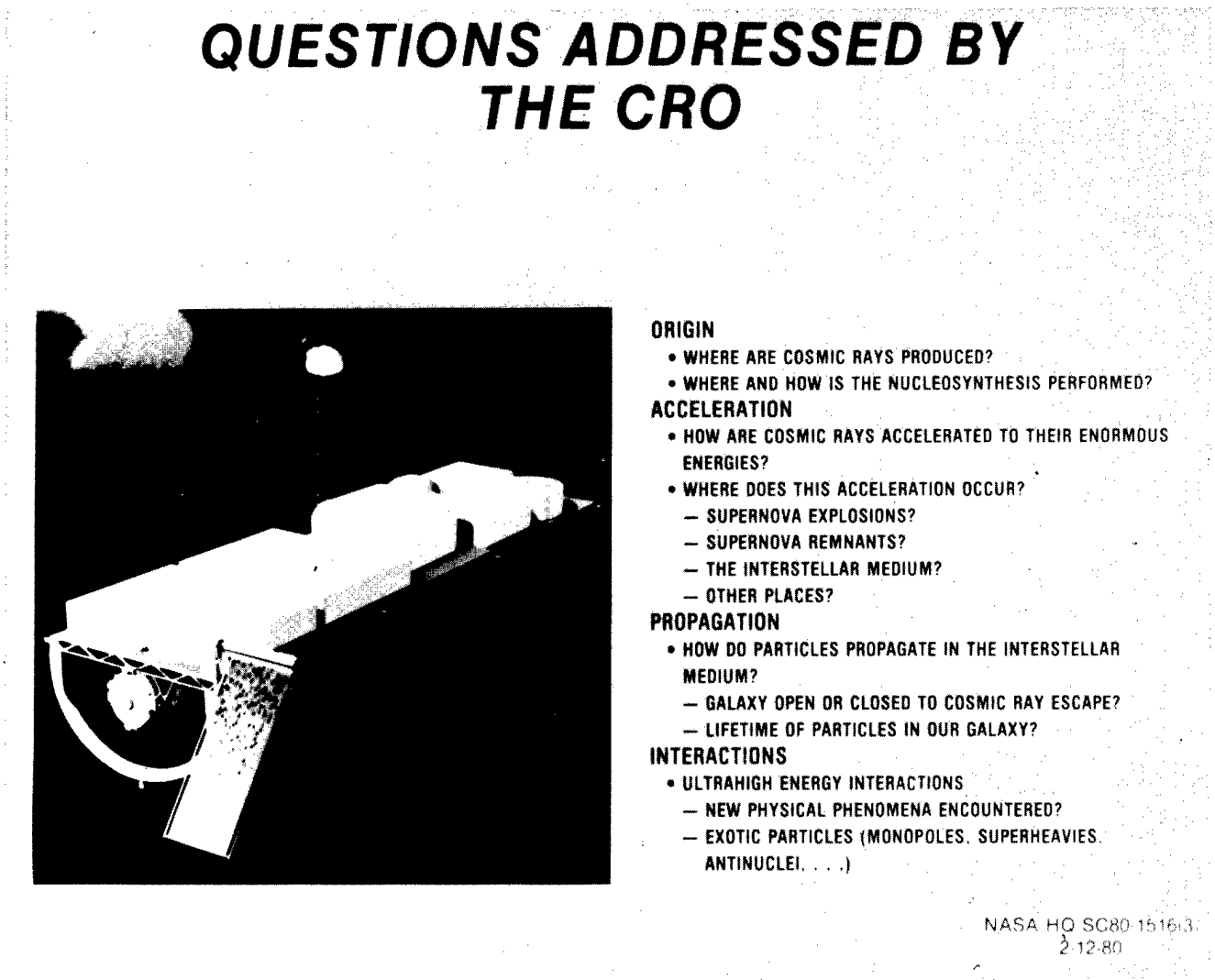


Figure 13. An artist's concept of the Cosmic Ray Observatory.

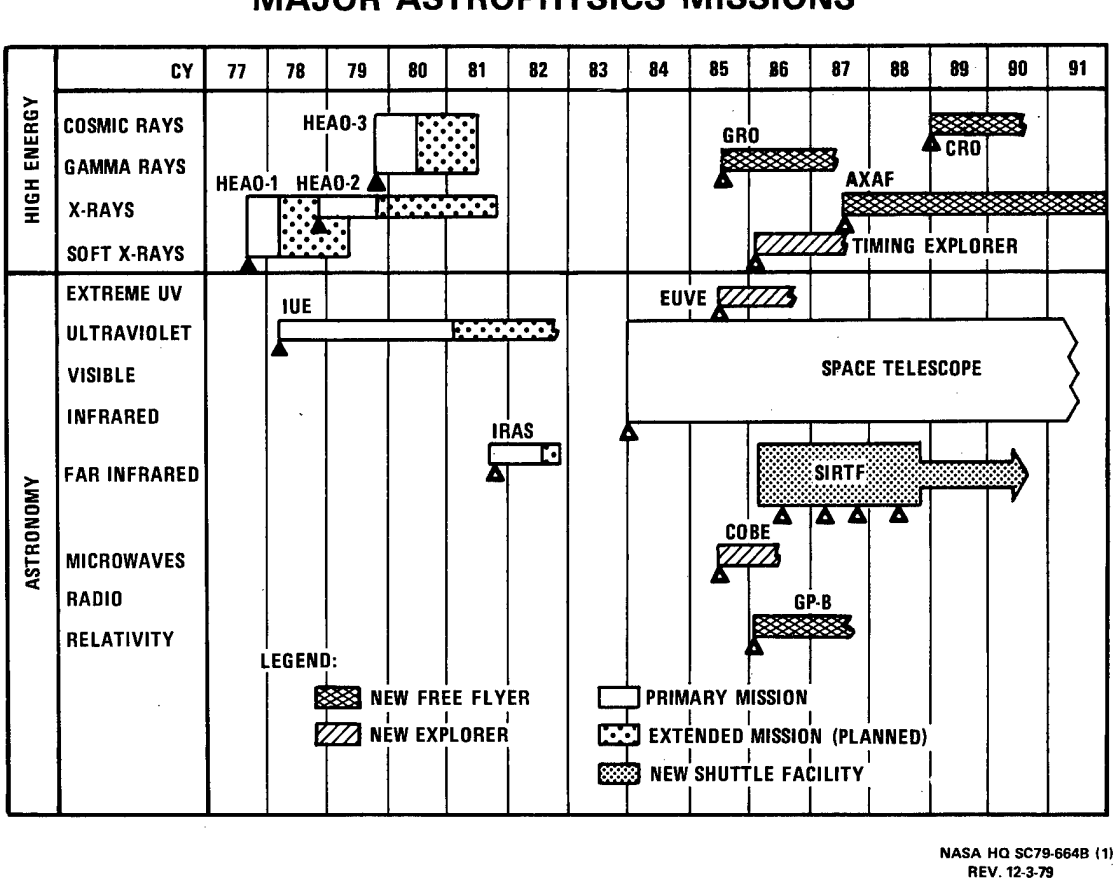


Figure 14. A possible flight schedule for missions in the most recent Astrophysics Five-Year Plan. The sequence shown assumes a new start for the Gamma Ray Observatory in Fiscal Year (FY) 1981, for Gravity Probe-B in FY 1982, for the Advanced X-Ray Astrophysics Facility in FY 1983, and for the Cosmic Ray Observatory in FY 1985.

the 1980s and into the 1990s. In X-ray astronomy, for example, we have already identified a need for two observatory-support type missions that will complement AXAF's capabilities.

X-Ray Observatory. This mission will carry a complement of X-ray instruments to make special measurements that do not require focusing optics and that systems such as AXAF cannot accommodate. Examples include iron line spectrometers, X-ray polarimeters, and instruments for studying energy ranges ($E < 0.4$ keV and $E > 10$ keV) which cannot be studied using AXAF.

Large Area Modular Array (LAMAR) X-Ray Telescope. LAMAR (Figure 15) will be an X-ray telescope having a very large collecting area and capable of conducting a very deep survey to establish the distribution of extragalactic X-ray objects. The distinguishing technical feature of LAMAR is its use of individual small modules (each with its own individual detector) which are then assembled to form a large X-ray collector (Figure 16). Both HEAO-2 and AXAF have sensitivities limited by their collecting areas of about 150 and 600 cm^2, respectively. It is envisioned that a collecting area as large as 10,000 cm^2 can be built up using the LAMAR approach. This will be done at the expense of angular resolution which will be about 1 arc minute. For reference purposes, the AXAF resolution goal is 0.5 arc seconds. Thus, LAMAR will be able to perform studies that do not require precise angular resolution but do require great sensitivity either because the sources are intrinsically faint or because they vary rapidly in intensity. Because of the novelty of the technical approach, development of competing LAMAR concepts for flight test is now being carried out as part of the Spacelab program.

We are also considering ways to use space to develop new observing capabilities or to dramatically extend existing ones. Based on previous astronomical history, the tools needed to address the astronomical problems of the 1990s and beyond will undoubtedly have to provide for increased angular resolution and larger apertures. It is highly likely that by the 1990s the ability to study many frontier astronomical problems will be limited by the weakness of the radiation received from intrinsically faint and extremely distant sources. As we consider the way new capabilities may develop in time, a number of approaches to this next generation of space observatories appear possible.

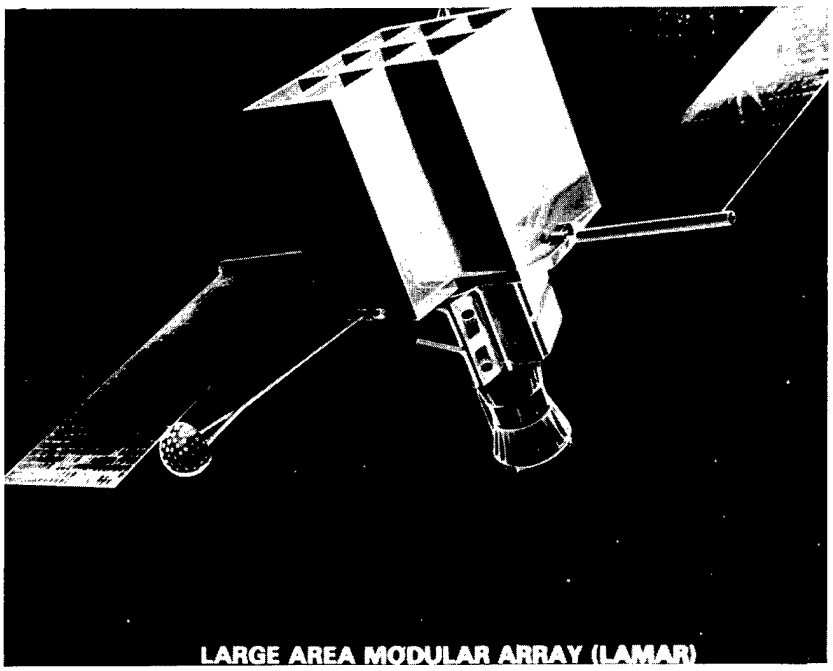


Figure 15. An artist's concept of the LAMAR X-ray Telescope.

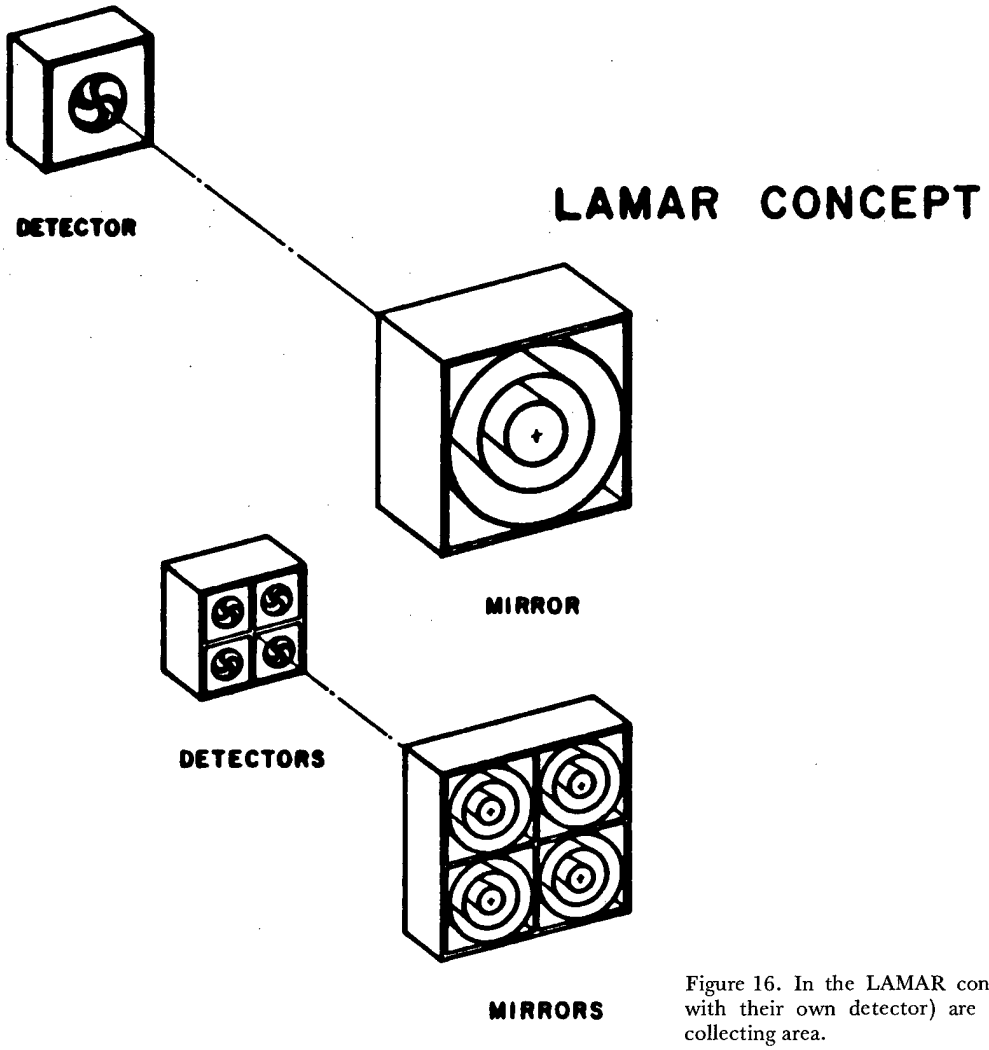


Figure 16. In the LAMAR concept individual small modules (each with their own detector) are combined to build up a very large collecting area.

Very Long Baseline Radio Interferometry. Very high angular resolution can be achieved in radio astronomy through an extension into space of the ground-based technique of Very Long Baseline Interferometry (VLBI). Use of an orbiting radio antenna 10 to 30 meters in diameter in conjunction with existing ground-based radio astronomy facilities (Figure 17) or other space-based antennas will permit production of detailed radio pictures of unprecedented angular resolution (10^{-3} to 10^{-5} arc seconds). Radio maps with that degree of detail are needed for the study of the physics of quasars and galactic cores, of interstellar masers and the dynamics of star formation. Use of an antenna in earth orbit will remove ambiguities that seriously affect interpretation of radio maps, particularly for complex sources. This gain comes about in the following way. Radio maps are generated by cross-correlating signals from the two antennas observing the source. The cross-correlation of a single simultaneous observation from the two antennas provides one component of a two-dimensional spatial Fourier Transform of the brightness distribution. To synthesize an unambiguous map, a full set of spatial Fourier components is required. Observations from many baseline orientations of the two antennas are, therefore, needed. A comparison of the coverage of the Fourier plane with Earth-based VLBI antennas and an orbiting VLBI terminal is shown in Figure 18. The advantage of space is obvious. Eventually use of two or more antennas in space and eccentric (or even interplanetary) orbits may also permit the use of significantly longer baselines (leading to higher resolution) than can be used on the Earth. For early applications at wavelengths of several centimeters or longer, a relatively simple mesh deployable antenna with a surface accuracy of 1 millimeter may be suitable. Preliminary tests of VLBI in space may be carried out with small antennas to be flown on the Space Shuttle.

Large flux collectors for infrared and submillimeter astronomy. At shorter wavelengths, more sophisticated technology will be required to generate large collecting surfaces with the more stringent surface accuracies which are needed. Preliminary studies of a 10-30 meter diameter aperture ambient temperature, deployable telescope for use in the infrared and submillimeter regions of the spectrum (Figure 19) are currently being carried out. Both regions are largely obscured by the Earth's atmosphere with image quality being set by atmospheric turbulence. Depending upon the exact surface accuracy and diameter achievable, such a telescope could be used for spatial mapping with angular resolution of order 0.1 arc-second at a wavelength of 10 microns, 1 arc-second at 100 microns, and 10 arc-seconds at 1 millimeter. A factor 4-12 better resolution than the Space Telescope in the wavelength region 10-1000 microns appears possible, permitting detailed studies of the structure of many important astronomical objects such as distant infrared galaxies, and regions of star formation and circumstellar shells in our own galaxy. The increased collecting area (in the range of 16-144 times that of Space Telescope) is important for carrying out very high resolution spectroscopic measurements of cool objects in our own and other galaxies. On-orbit assembly of the telescope (Figure 20) is also a possible approach for building such a large flux collector.

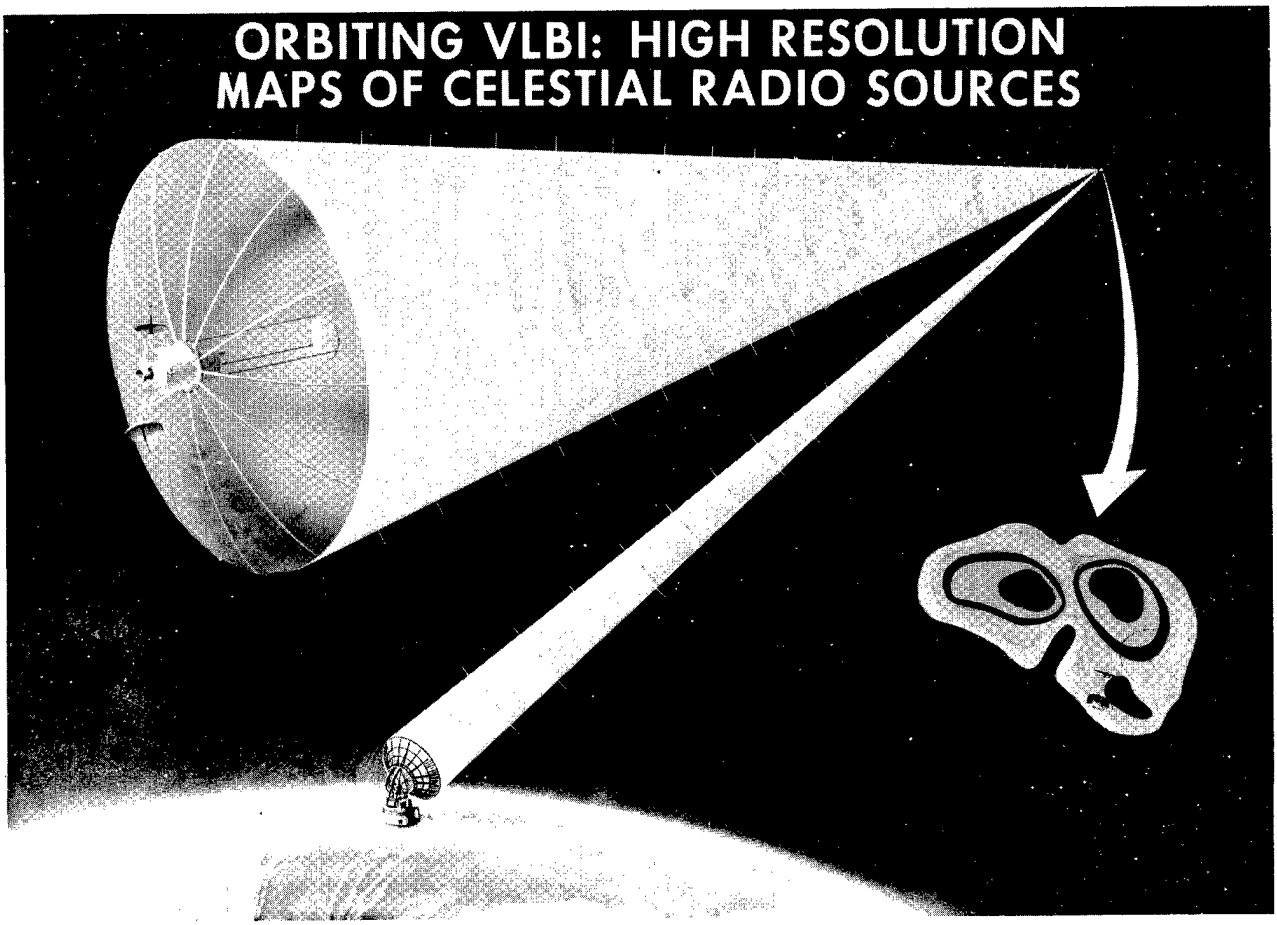


Figure 17. A VLBI receiver pair with one Earth-based terminal and one orbiting terminal.

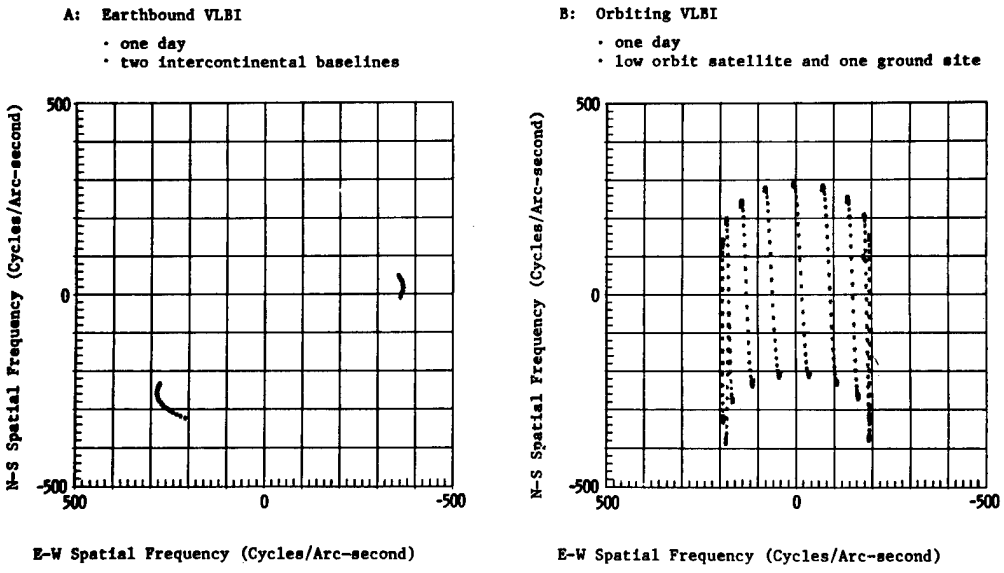


Figure 18. Comparison of the mapping capability of earth-based and orbiting VLBI. Each data point represents the measurement of a spatial Fourier component of some radio source brightness distribution.

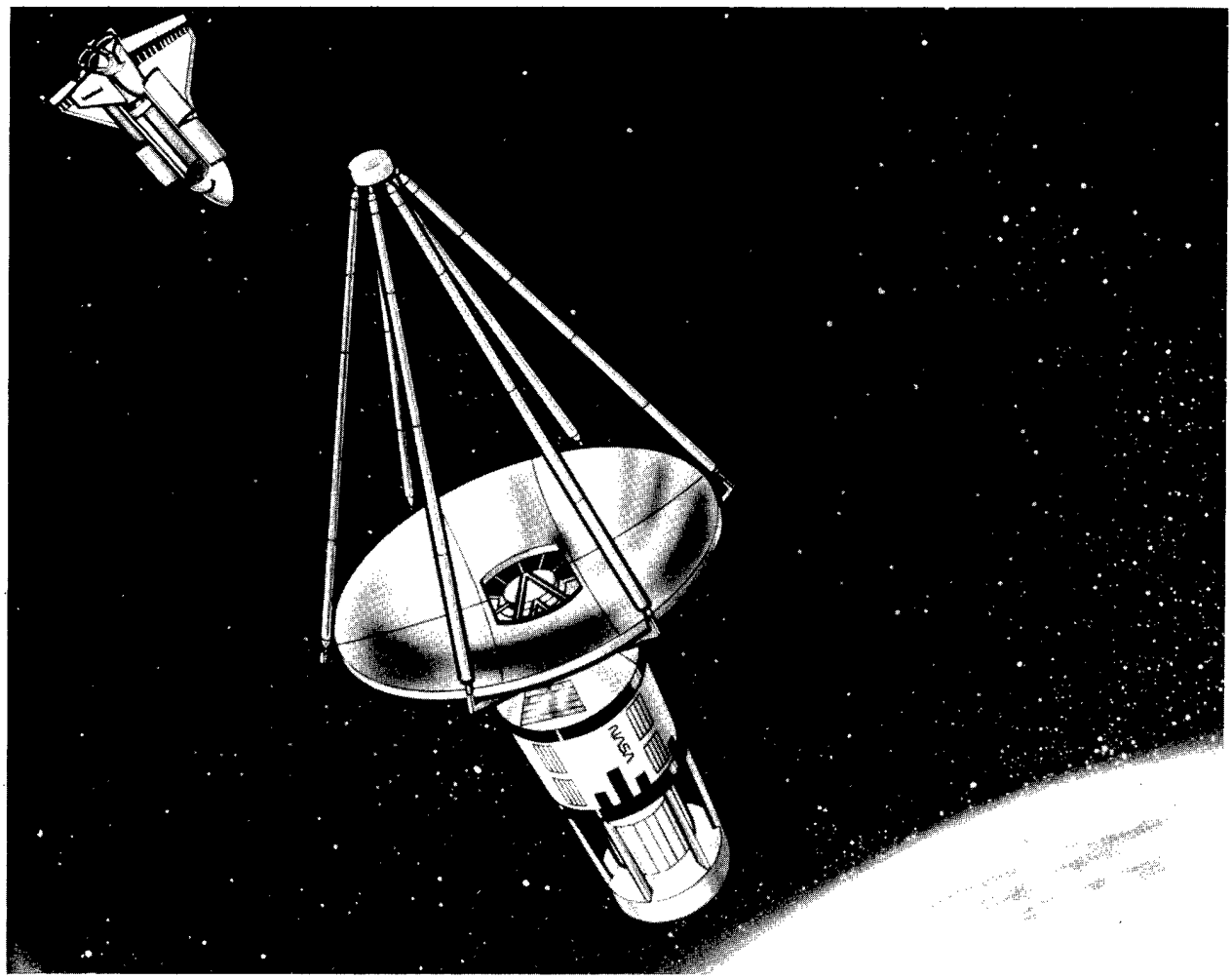


Figure 19. An artist's concept of a Large Ambient Deployable Infrared Telescope. Such a large flux collector would also be useful for submillimeter and millimeter wave astronomy.

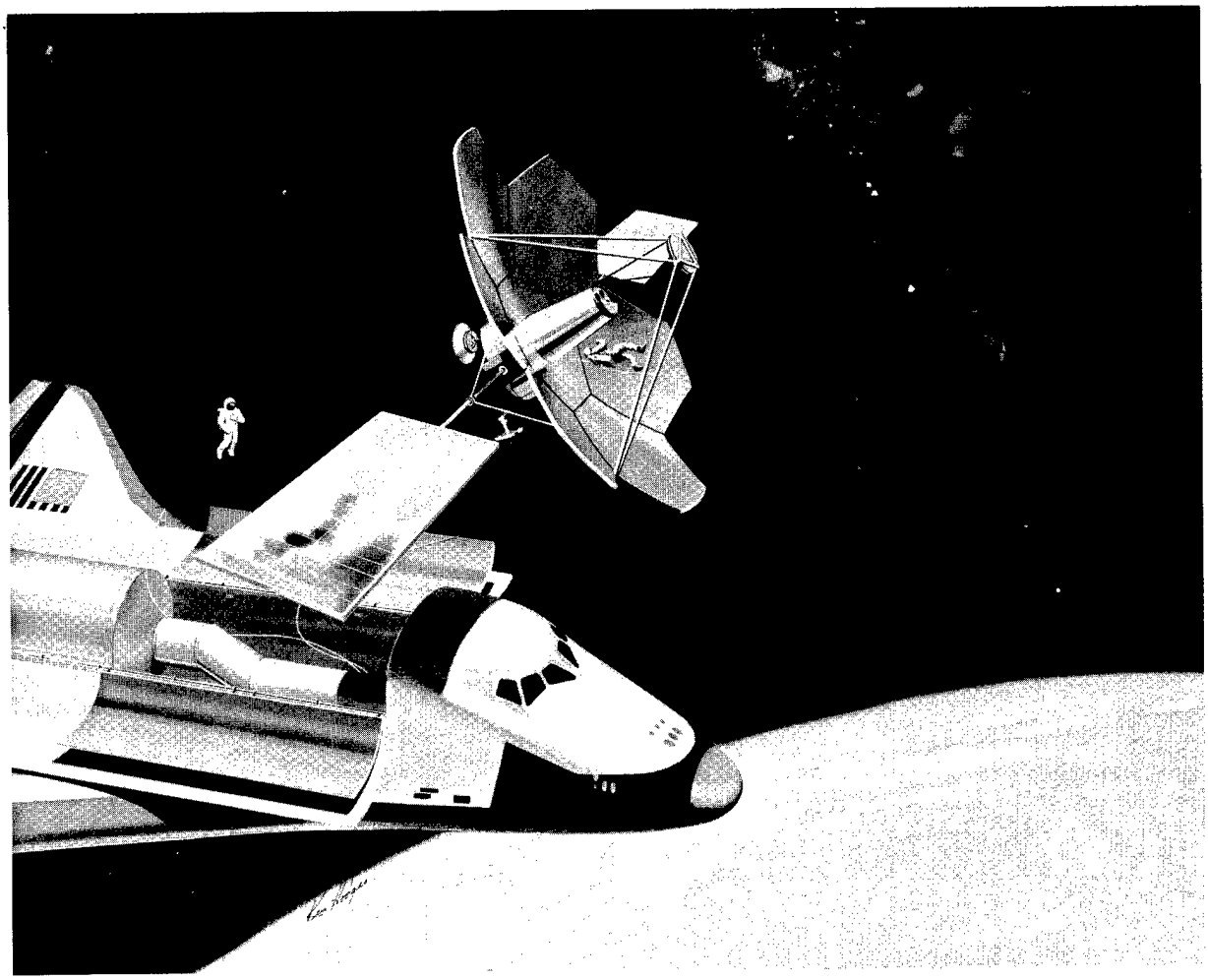


Figure 20. Astronauts may be used to assemble large flux collectors in space.

Arrays for use in UV/Optical Astronomy. As we look toward the end of the century, the usefulness of the Space Telescope will ultimately be limited by the fact that less than one photon per second will be collected by the 2.4 meter diameter mirror for many of the objects of greatest astronomical interest. Because of the extreme tolerances required in the UV/Optical region of the spectrum, the deployable or erectable approaches which appear attractive at infrared or microwave wavelengths seem less appropriate for near diffraction limited visible light imaging, and a variety of other design concepts are being studied. In one approach, a large (6-10 meter aperture) two-mirror UV/Optical telescope could be put together in orbit using preassembled segmented optical components. An aperture in this size range is required for a ten-fold gain in sensitivity over the Space Telescope. Such a mirror could be carried in a modified Space Shuttle external tank, with the complete Observatory being assembled in Space (Figure 21). Even higher angular resolutions and larger collecting areas appear possible using advanced concepts involving arrays of mirrors or mirror segments. Telescope modules may be combined to form a large optical array (Figure 22). Such a system might start out as a simple 2 or 4 element linear interferometer with a 10 meter baseline and, with the addition of more modules, eventually evolve into a true imaging system. The ultimate goal would be to achieve an imaging quality of 10^{-3} arc-seconds and an astrometric accuracy of 10^{-6} arc-seconds, several order of magnitude gains over the capabilities of the Space Telescope. In particular, such an astrometric accuracy would permit detection of an earth-mass planet around a solar-type star out to a distance of 30 light years. Space-based interferometers may also be useful at infrared wavelengths (Figure 23). Even more radical concepts (Figure 24) may be used to build up light collecting areas equivalent to a 70-meter aperture telescope. All of these possibilities will depend in a fundamental way upon having the capability to assemble large structures in space and the use of advanced optical control technology to maintain baselines and alignments and to achieve high accuracy pointing of large structures.

Gravity waves: a possible new astronomy. Large, controlled structures in space (Figure 25) may also be used to try to detect gravity waves. Relativistic theories of gravitation predict that gravitational radiation is generated by any non-spherical dynamically changing system of masses. Possible sources of gravity waves include pulsars, short-period binary stars containing compact objects, and massive rotating stars that are collapsing to form black holes. Different types of sources are expected to emit gravitational radiation of different characteristic frequencies (Figure 26). In particular, the very low frequency (VLF) gravity waves may be detectable using a large interferometer. Such an instrument might consist of four 1000 kilogram masses suspended at the ends of two crossed arms each one kilometer long. The masses would serve as mirror mounts in a multipass Michelon interferometer. Measurements of the relative displacements of the masses would be used to detect gravity waves. Possibilities such as this raise the fascinating prospect of a new era of observational gravitational astronomy.

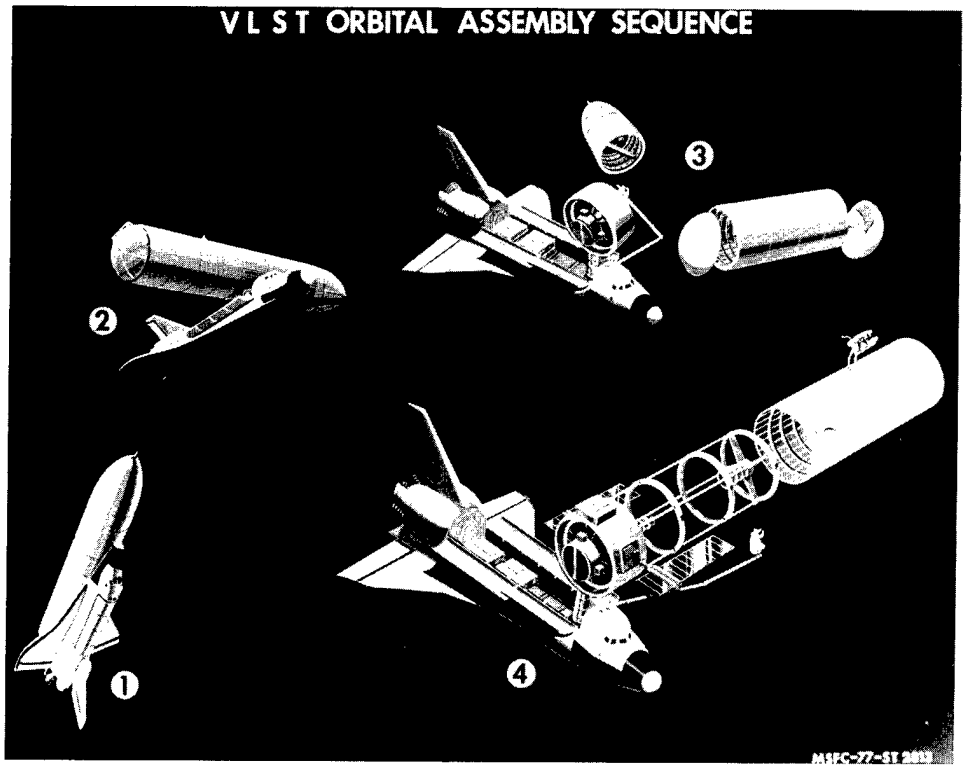


Figure 21. Possible assembly sequence for a Very Large Space Telescope. Various components of the telescope may be brought up on successive Shuttle flights.

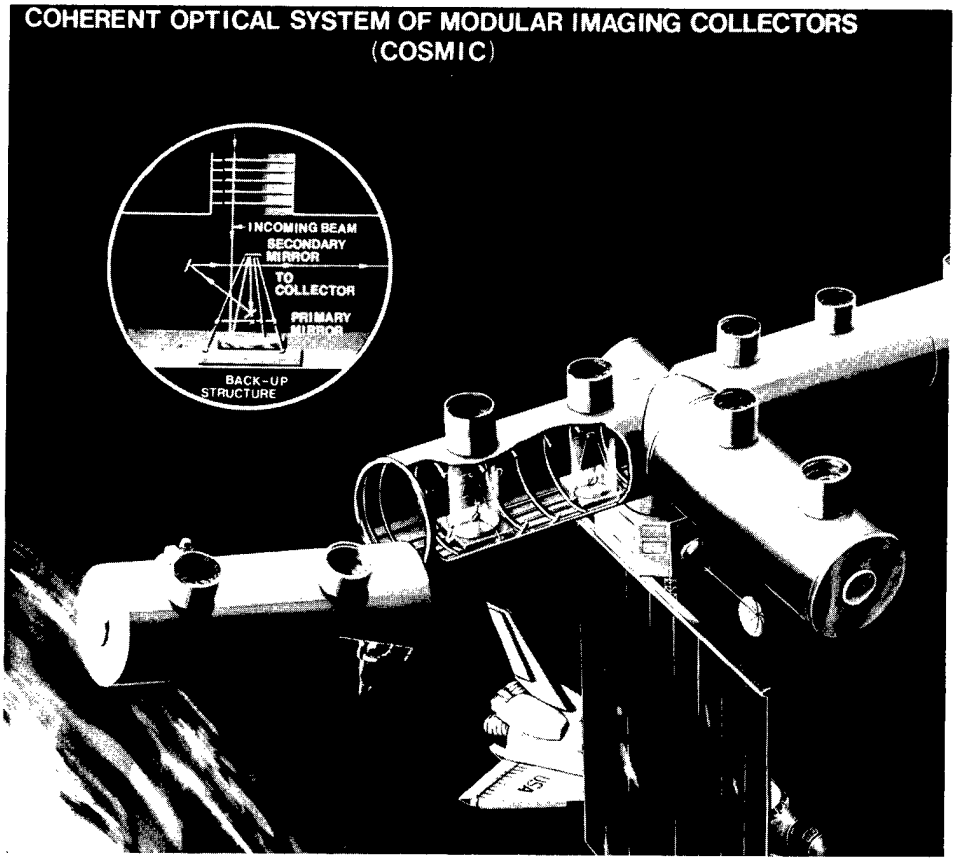


Figure 22. Assembly of a large coherent array of optical telescopes. Such arrays may be used to achieve very high angular resolution as well as to build up a large collecting area.

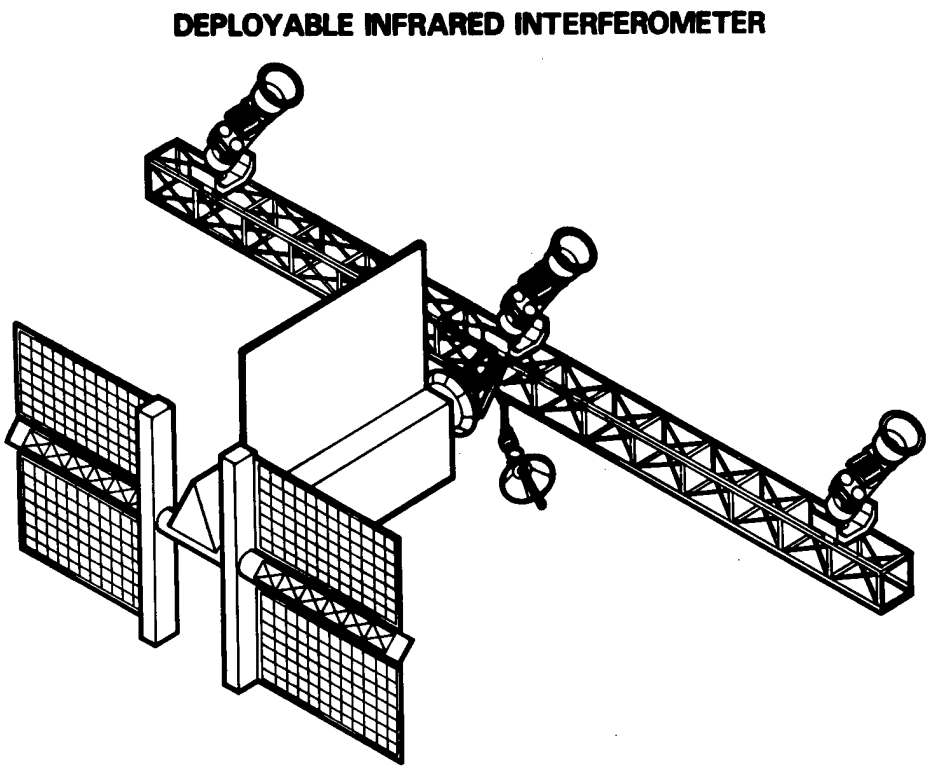


Figure 23. An artist's concept of an infrared interferometer. Such an interferometer might be composed of a number of cryogenically cooled telescopes deployed along a structural beam and attached to a free flying power module.

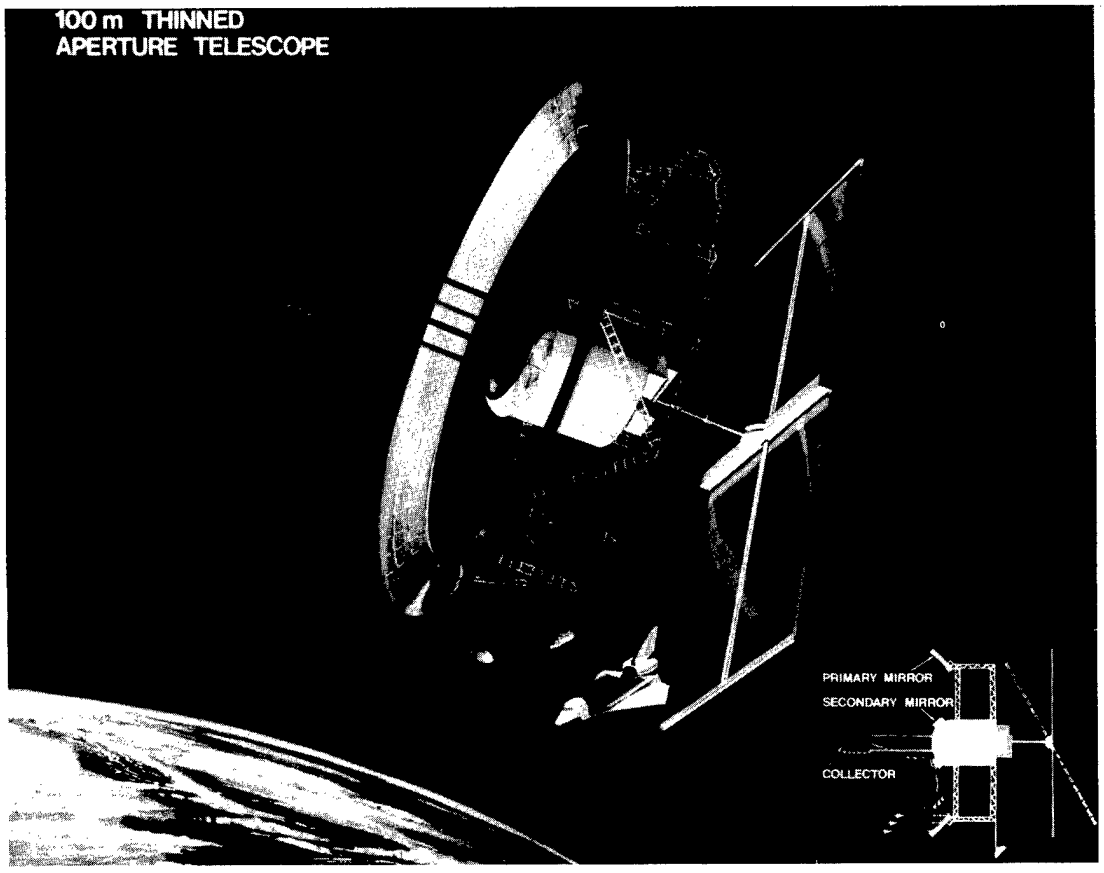


Figure 24. An artist's concept of a very large annular telescope. Use of novel optical designs could lead to a relatively compact configuration.

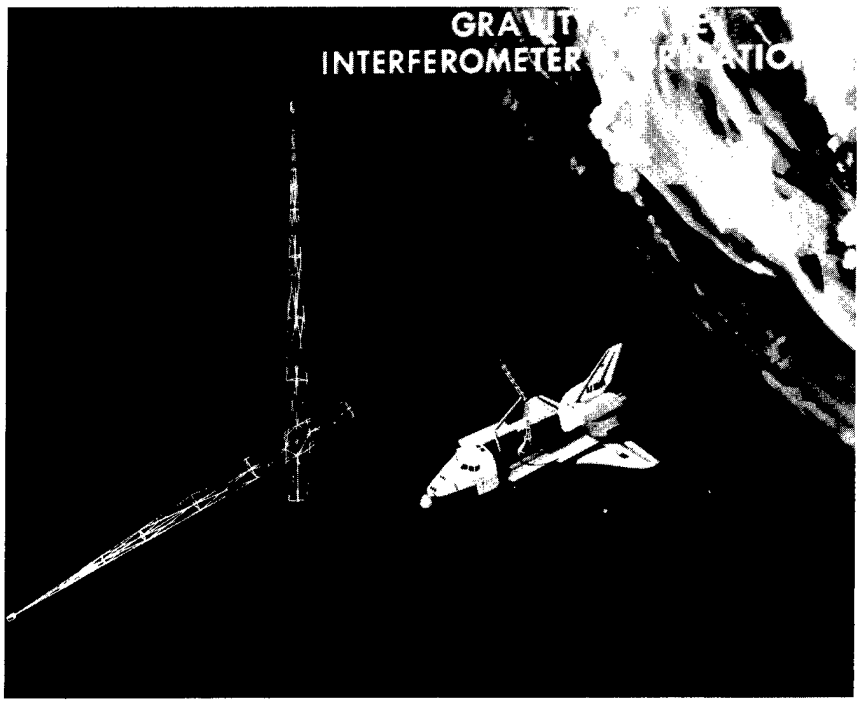


Figure 25. Assembly in space of a large gravity wave interferometer. The structural members may actually be fabricated in Space using a shuttle-based beam building machine.

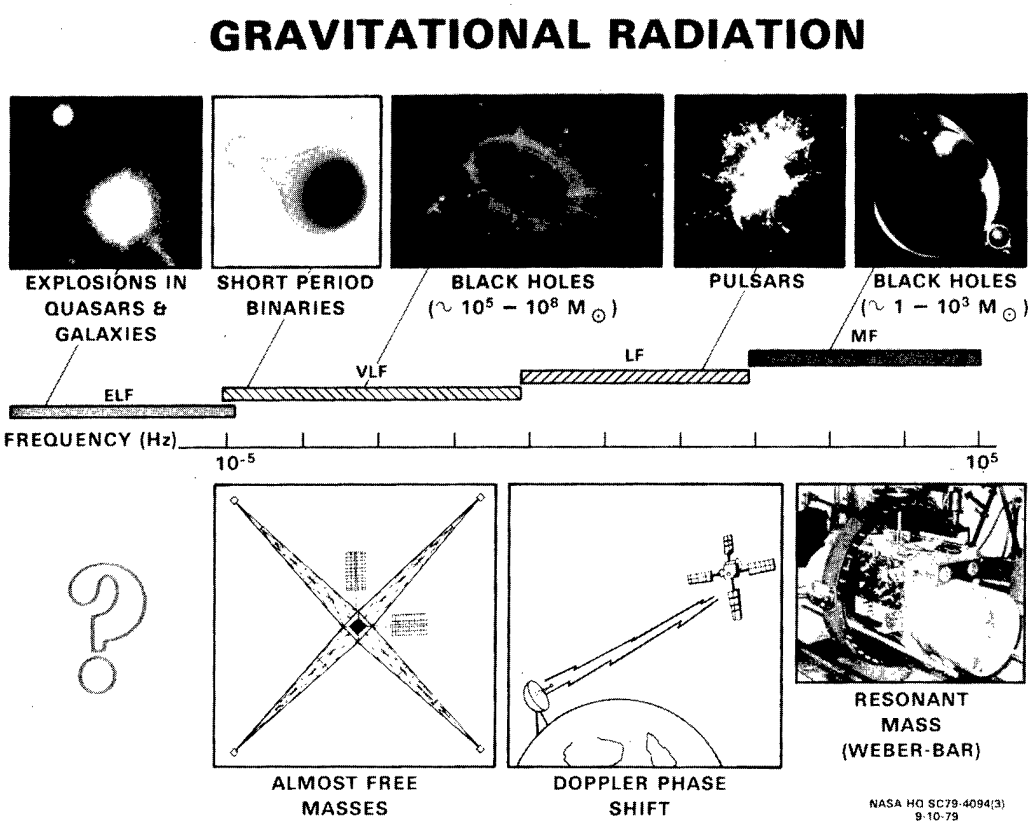


Figure 26. Different sources produce gravitational radiation having different characteristic frequencies. Space techniques may be most useful for detection of Low Frequency (LF) and Very Low Frequency (VLF) gravity waves.

Concluding remarks

The ultimate goal of astrophysics is to understand the origin and evolution of the Universe. This is a bold (and some might even say foolhardy) undertaking. It has been remarked that the Universe is not only stranger than we imagine, it is stranger than we can imagine. Access to space has provided and will continue to provide the astrophysicist with a whole new set of tools and techniques for addressing fundamental questions. A veritable zoo of bizarre astrophysical "creatures" has been discovered during the past twenty years, and the space program has played a central role in these discoveries. The future offers the prospect of an even richer range of possiblities. With the development of suitable technology the next generations of space experiments can provide the next steps forward in achieving an understanding of just how strange the Universe is.

Acknowledgements

I would like to thank Dr. Samuel H. Morgan, Jr. for his invaluable assistance in the preparation of graphical material and in helping with the layout of this paper.

Appendix

Dictionary of Acronyms

AXAF	–	Advanced X-Ray Astrophysics Facility
COBE	–	Cosmic Background Explorer
COSMIC	–	Coherent Optical System of Modular Imaging Collectors
CRO	–	Cosmic Ray Observatory
EUVE	–	Extreme Ultraviolet Explorer
GP-B	–	Gravity Probe B
HEAO	–	High Energy Astronomy Observatory
IRAS	–	Infrared Astronomical Satellite
IUE	–	International Ultraviolet Explorer
LAMAR	–	Large Area Modular Array X-Ray Telescope
OAO	–	Orbiting Astronomical Observatory
RAE	–	Radio Astronomy Explorer
SAS	–	Small Astronomy Satellite
SIRTF	–	Shuttle Infrared Telescope Facility
ST	–	Space Telescope
VLBI	–	Very Long Baseline Interferometry
VLST	–	Very Large Space Telescope
XTE	–	X-Ray Timing Explorer

Advanced x-ray astrophysics facility (AXAF)

Martin C. Weisskopf
Space Sciences Laboratory, NASA/Marshall Space Flight Center
Marshall Space Flight Center, Alabama 35812

Abstract

The Advanced X-Ray Astrophysics Facility (AXAF) will be a national observatory designed for the observation of galactic and extragalactic X-ray sources. The observatory is currently planned to be launched by the Space Shuttle, maintainable in orbit and retrievable, if necessary, during its 10 to 15 years of operation. The heart of the AXAF is an X-ray telescope made up of six Wolter type I mirrors, with the diameter of the outermost mirror 1.2 m. The focal length will be 10 m, and the AXAF will allow for interchanging and (in-orbit) replacing of focal plane instruments. The optics are being designed to provide a spatial resolution of 0.5 arc second over a several arc-minute field and somewhat reduced angular resolution over the entire 1-degree field of view. The energy bandwidth will be 0.1 to 8 keV. These design goals place severe requirements on the materials, tolerances, construction, and alignment of the telescope. These will be discussed and compared to previous work in this area.

Introduction

Our knowledge of the universe is obtained by studying the energy emitted by astrophysical objects. For most of history this has meant the study of visible light, and man's fascination with the heavens is known to date back as far as the records of the moon's phases scratched on chunks of bone more than 30,000 years ago. The "modern" era of astronomical observations began in 1610 when Galileo used the optical telescope to discover the moons of Jupiter. Since then, astronomers have sought ways to improve their view of the sky. Visible light, however, is only one of the many types of electromagnetic radiation emitted by celestial objects. Figure 1 identifies the other regions of the electromagnetic spectrum and illustrates why astronomers have only recently been able to view the universe in other than the visible; that is, many types of radiation cannot penetrate the earth's atmosphere. The field of X-ray astronomy, in particular, is a product of the space age which now allows scientists to perform astrophysical observations above the filtering effects of the earth's atmosphere.

X-rays, at the high energy end of the electromagnetic spectrum, play a unique role in providing detailed insight into the "how" and "why" of the physical processes that govern the universe. This follows in part because of the high energies required to create X-rays and in part because of the strange and fascinating objects which have been discovered to emit them. These objects including spinning neutron stars (so dense that a piece of neutron star material the size of a raindrop weighs as much as the combined weight of 3 billion people), supernova remnants (the remains of stellar explosions), and black holes (objects whose mass and radius have evolved so much that they totally collapse, leaving a "hole" in space from which not even light can escape). However, these objects are far from being just interesting oddities; on the contrary, they play a major role in the history of the universe. In addition, stars, normal galaxies, exploding galaxies, clusters of galaxies, and the mysterious quasars number among the X-ray emitters.

Because of X-ray observations as well as those in other regions of the spectrum, especially the radio, astronomers have overturned the age-old concept of a constant static universe in favor of one which is dominated by change and violent phenomena. X-ray astronomy has played, and will continue to play, an important role in furthering our understanding of the universe because X-rays are always produced wherever and whenever such phenomena take place.

The advanced x-ray astrophysics facility

The Advanced X-Ray Astrophysics Facility (AXAF), currently under study, is conceived of as the next major step in the field of X-ray astronomy. The AXAF will be an X-ray observatory built around a large-area, high-resolution grazing incidence X-ray telescope and designed to operate in space as a national facility for 10 to 15 years. An artist's conception of the AXAF in orbit is shown in Figure 2. Also shown is the Space Shuttle which will be used to place the AXAF above the earth's atmosphere. Data from the AXAF's scientific instruments will be transmitted to the ground via NASA's Tracking and Data Relay Satellite System (TDRSS).

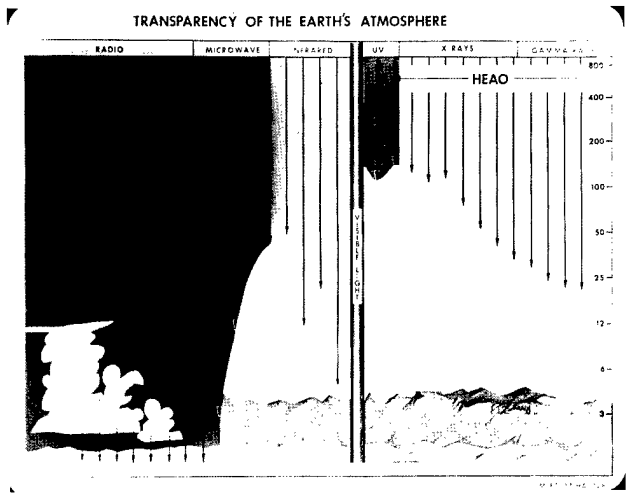


Figure 1. Transparency of the earth's atmosphere to electromagnetic radiation as a function of energy. The arrow tips end at the highest altitude to which a particular wavelength penetrates.

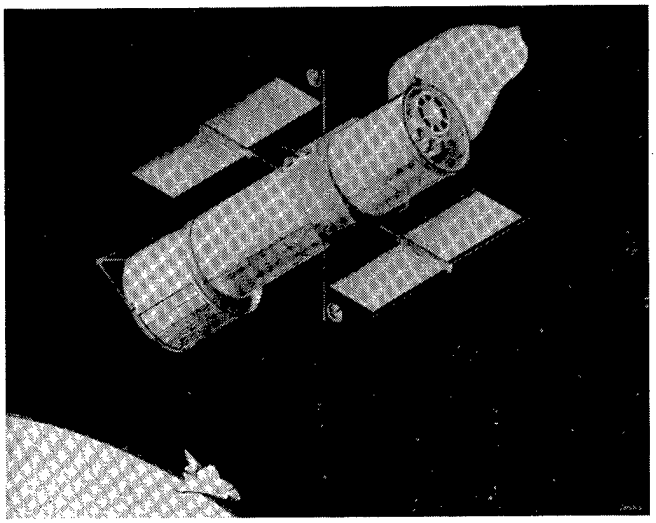


Figure 2. An artist's conception of the AXAF in orbit.

The long lifetime of AXAF will provide scientists with a facility capable not only of performing the observations now known to be necessary because of previous discoveries and the questions raised by them, but also to expand X-ray observations into uncharted waters. The AXAF will make use of the versatility offered by the Space Transportation System which allows for the refurbishment and replacement of instruments in orbit. Thus one can take maximal advantage of the advances in detector technology which will take place throughout the lifetime of the observatory.

The scientific goals and the technical foundation of the AXAF rest heavily on the spectacular success of the HEAO-2 (Einstein) X-ray Observatory which was launched in November 1978 and will reenter the atmosphere in early 1981. Furthermore, the Einstein Observatory, through its Guest Investigator Program, is the most heavily utilized satellite that NASA has ever flown. This is testimony to the enthusiastic support from within the astrophysical community for X-ray astronomy in general and the AXAF in particular. The AXAF, however, represents a significant improvement over the Einstein Observatory in all respects. Two unique features of the AXAF have already been mentioned: namely, its long lifetime and the capability for the replacement and refurbishment of instrumentation. In addition, and of equal importance, the AXAF will have four other major improvements over the Einstein Observatory: higher angular resolution, broader energy coverage, larger effective area, and the capability of a full complement of instrumentation to study all the properties of electromagnetic radiation in the X-region of the spectrum -- spatial resolution, energy resolution, time resolution and sensitivity to polarization. Thus, AXAF will blend a significantly increased capability and sensitivity for X-ray observations in wavelength regions where the Einstein Observatory obtained preliminary results to guide the way, with an extension of X-ray imaging to higher energies where new results await us. With an average sensitivity of fifty to one hundred times that of any previous X-ray experiments and a lifetime five times longer than any previous X-ray mission, the AXAF will allow, for the first time, the detailed study of all the known categories of X-ray-emitting objects. This capability places X-ray astronomy in the position to complement and coordinate with the other major astronomical observatories of the 1980's, the Space Telescope (ST) for visible light and the Very Large Array (VLA) for radio observations. Together with AXAF the observatories will provide unique and powerful tools for man's exploration of the universe through the remainder of the twentieth century.

The high-resolution mirror assembly

The heart of the AXAF is the grazing incidence X-ray telescope. The baseline design is shown in Figure 3 and consists of a Wolter type I mirror assembly with six nested paraboloid-hyperboloid pairs. The inner diameter of the outermost pair is 1.2 m and the focal length is 10 m. The length of each segment (paraboloid and hyperboloid separately is 84 cm. Adjacent optical surfaces are separated by distances on the order of centimeters to avoid vignetting

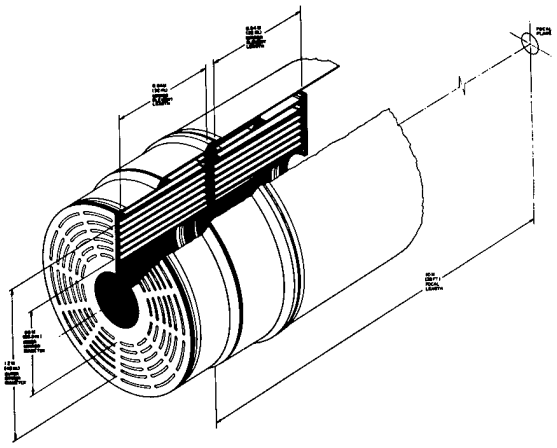


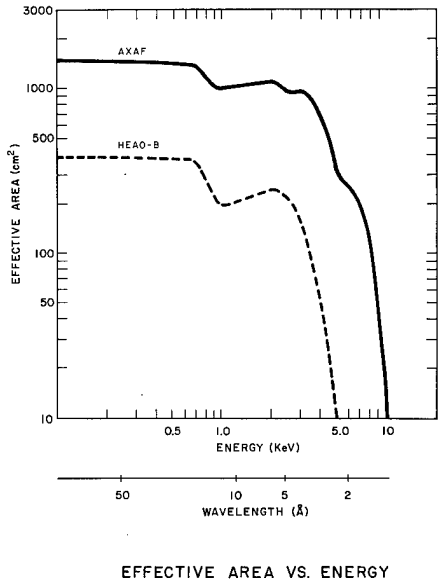
Figure 3. AXAF grazing incidence telescope conceptual configuration.

of X-rays within 20 minutes of arc from the optical axis and also to allow sufficient structural thickness to the mirror walls.

The baseline design has evolved from the interaction between scientific requirements and engineering constraints. It has been enhanced as a result of the experience gained in the design, fabrication, assembly, test, and performance of the Einstein telescope. The AXAF telescope characteristics are summarized in Table 1, where they are also compared to those of the Einstein Observatory. There are several fundamental differences between the two telescopes, and these are only partially apparent in Table 1. First of all, the AXAF has a larger geometric collecting area, making it far more efficient than the Einstein telescope. Furthermore, the range of grazing angles allowed by the AXAF design runs from 27 minutes of arc for the inner mirror pair to 51 arc minutes for the 1.2 m outer mirrors. Since the maximum X-ray energy that can be reflected is inversely proportional to the grazing angle, the smaller grazing angles of the AXAF allow the response of the telescope to extend to energies above 7 keV. The total on-axis effective collecting areas are compared in Figure 4. The off-axis response of the AXAF is shown in Figure 5.

Table 1. A Comparison of Telescope Parameters

	HEAO-2/Einstein	AXAF
No. of Elements	4 Nested Pairs	6 Nested Pairs
Outer Diameter	0.58 m	1.2 m
Focal Length	3.44 m	10.0 m
Geometric Area	460 cm^2	1700 cm^2
Inner Mirror Grazing Angle	0.68°	0.45°
Outer Mirror Grazing Angle	1.17°	0.85°
Resolution (FWHM)	4.0 arc sec	0.5 arc sec (Goal)
Field of View	1°	1°
Fraction of flux within a 1 arc sec diameter		
@ 1.0 keV	<3%	>90%
@ 2.5 keV	<3%	>60%
@ 5.0 keV	<0.5%	>35%
@ 7.0 keV	0	>20%



EFFECTIVE AREA VS. ENERGY

Figure 4. Effective area as a function of energy for on-axis radiation.

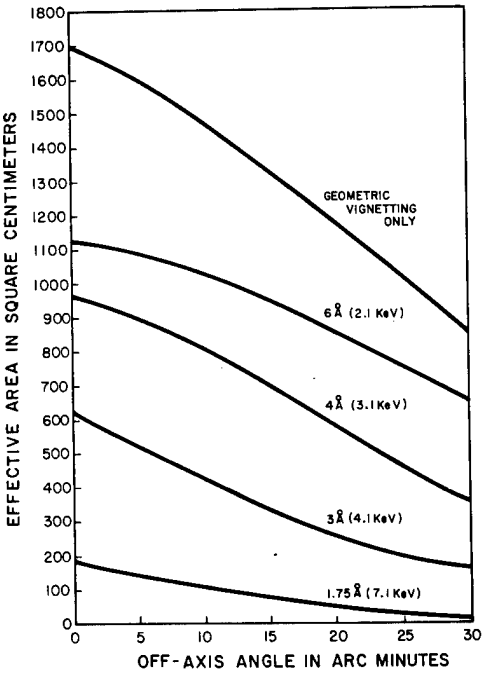


Figure 5. Effective area of the AXAF telescope as a function of off-axis angle at selected wavelengths (energies).

The extension of X-ray imaging beyond the 7 keV has important scientific consequences. In general, a new regime of the X-ray spectrum is opened for X-ray imaging. But, further, this allows for observations which include the important complex of lines of highly ionized iron between 6 and 7 keV. These lines are a dominant feature in the X-ray spectra of many astrophysical objects. The ability of the AXAF to perform X-ray imaging at these energies will thus provide an important diagnostic for probing the composition, structure, and general physical nature of a wide variety of systems.

The ability to extend the realm of X-ray imaging to high energies is partially a consequence of the longer focal length of the AXAF which, for a given diameter, allows for more significant geometric areas at lower angles of incidence. The long focal length of the AXAF has a second important consequence, namely a large plate scale (50 μm per second of arc, which is a factor of three greater than that of the Einstein Observatory. This automatically yields a threefold increase in angular resolution for even the HEAO generation of focal plane instrumentation.

The most significant difference between the AXAF and its prototype is the imaging quality. The AXAF's enhanced spatial resolution and sensitivity (determined by the fraction of flux within the core of the image) are of the highest scientific consequence. Thus, for example, in the study of extragalactic objects, a large number of astrophysical problems cannot be attacked at the 4 arc second resolution level of the Einstein. These include the capability of resolving discrete X-ray sources in other galaxies. Such observations could only be achieved of M31 (Andromeda, our nearest companion galaxy) with, however, spectacular results (Figure 6). The AXAF with its effective area and angular resolution would allow one to extend such observations to much more distant galaxies.

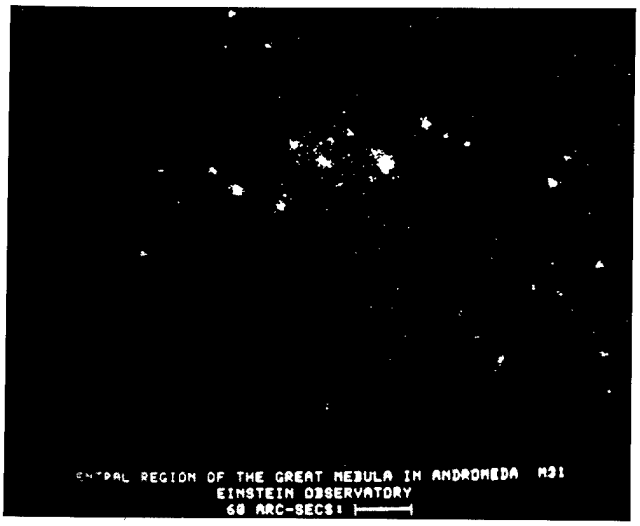


Figure 6. Computer reconstruction of an electronic photograph of the center of M31 taken with the high resolution imaging X-ray detector on the HEAO-2/ Einstein Observatory. M31 (Andromeda) is a nearby galaxy. More than 70 X-ray sources have been discovered in this galaxy, thus, equalling, in a single observation, the results of early maps of our own galaxy and providing a wealth of new astrophysical information. (Picture courtesy of Prof. R. Giacconi, Harvard-Smithsonian Center for Astrophysics).

The imaging performance of the AXAF will be determined, not by diffraction, but by the geometrical figure, alignment, and surface finish of the mirrors. The figure and alignment determine the full width half maximum (FWHM) of the point spread function (PSF). The surface finish, primarily microscopic surface roughness, determines the fraction of the reflected flux that remains within any given radius about the center of an X-ray image. The design tolerances of the AXAF are currently such that figure errors will be kept at a level that will provide a PSF with a FWHM of 0.5 second of arc. Since the FWHM of an X-ray telescope varies slowly with angle of incidence, high contrast features can be resolved out to the edge of the field of view. The fraction of the imaged flux that lies within the FWHM is, however, a strong function of energy (as shown in Figure 7) and decreases rapidly with the angle of incidence. Another important feature of X-ray telescopes is that the focal surface, which provides optimum imaging, is not a plane. Figure 8 shows the focal surface obtained by minimizing the rms blur circle diameter in a geometric ray tracing routine. In this calculation a reflectivity of unity was assumed, and the resulting focal surface can be approximated by a spherical shape with a radius of 96 mm and a pole at the on-axis gaussian focus. Thus, a narrow field, high spatial resolution detector might be best placed, not at the gaussian focus, but 0.15 mm towards the mirror assembly in order to achieve a more flat response over more than 1.5 minutes of arc, as shown in Figure 9.

The AXAF telescope performance discussed previously must be translated into fabrication tolerances on the mirror assembly. The method of specifying fabrication tolerances that we use is a direct outgrowth of the Einstein Observatory experience and is based on a procedure which converts the required X-ray performance into parameters that are amenable to measurement during fabrication and assembly. Figure 10 illustrates both the definition and current specifications of the AXAF fabrication tolerances. In Table 2 tolerances are compared with the corresponding Einstein requirements and with the values actually achieved during the Einstein fabrication and assembly. We note that in several cases the AXAF tolerances are no more stringent than they were for the Einstein. On the other hand, the most demanding of these tolerances -- that the variation in the difference between forward and aft radii as a function of azimuth [$\Delta(\Delta R)$] be less than $\pm$ 15 microinches, that the variation between the actual surface slope and the design slope be less than $\pm$ 0.25 microradian, and that the surface roughness be less than 15 Å rms -- will require the full advance of fabrication technology that has taken place in the intervening years since the HEAO-2 was constructed. There is no indication, however, based on our technology studies discussed later, that any of these requirements are beyond the state-of-the-art.

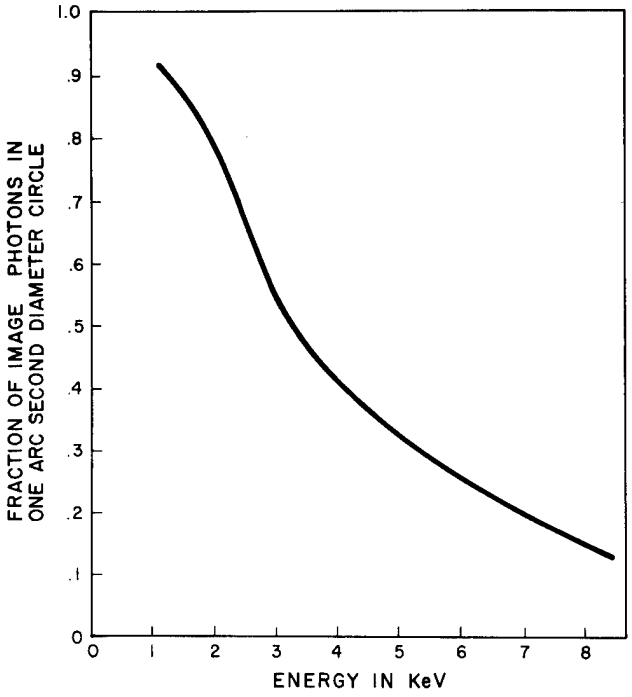


Figure 7. Fraction of reflected flux in the central (1 arc second diameter) image as a function of energy. The calculation was performed assuming a Beckmann model for the scattering distribution.

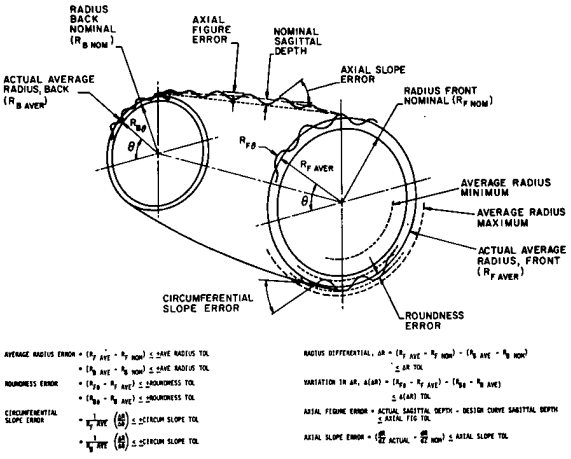


Figure 10. Schematic representation and definition of mirror element tolerances.

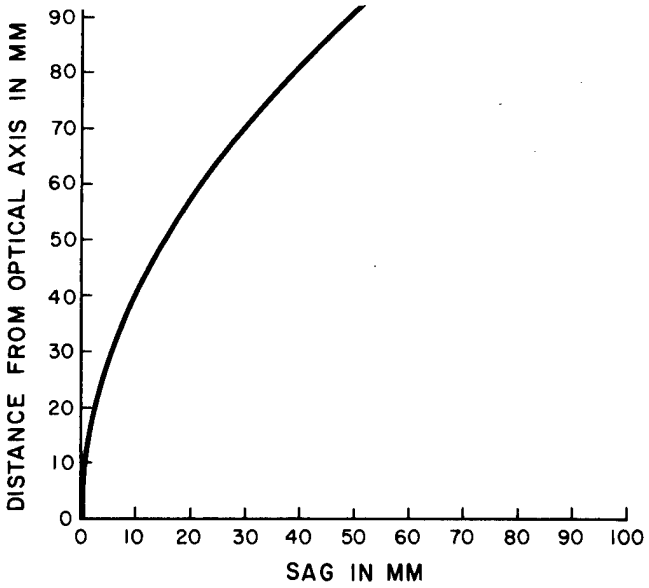


Figure 8. AXAF focal surface obtained by minimizing the rms blur circle diameter in a geometric ray tracing. A reflectivity of unity was used.

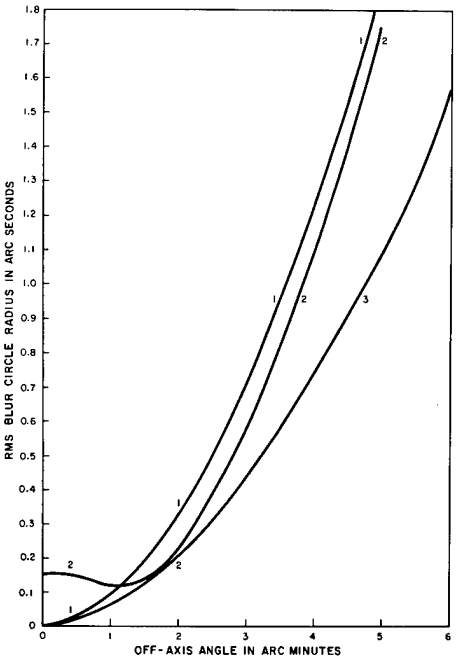


Figure 9. RMS blur circle radius versus off-axis angle at(1) Gaussian focal plane,(2) focal plane placed 0.15 mm towards the telescope,(3) "optimum" focal surface shown in Fig. 8.

Table 2. Preliminary AXAF Limiting Mirror Tolerances and Comparison with HEAO-2

Parameter	AXAF	HEAO-2	Achieved on HEAO-2
$\Delta(\Delta R)$	$\pm 15\times10^{-6}$ inch	$\pm 25\times10^{-6}$ inch	$\pm 22\times10^{-6}$ inch (avg)
Axial Slope	0.25×10^{-6} radians	$\pm 3\times10^{-6}$ radians	Yes
Circumferential Slope	$\pm 2\times10^{-6}$ radians	$\pm 25\times10^{-6}$ radians	4.7×10^{-6} radians (avg)
Coma/Mirror Pair	0.1 arc sec rms	Minimize	0.4 arc sec rms
Coincidence of Faci			
Lateral	0.0002 inch dia	0.0005 inch dia	0.0002 inch dia
Axial	0.0016 to 0.003 inch	0.002 inch	0.003 inch
Surface Roughness	15 Å rms	30 Å rms	14 to 25 Å rms

In addition to achieving the fabrication tolerances, means for in-process and possibly in situ measurements of figure and surface roughness will be desirable to reduce the set-up and turn-around time for these measurements during fabrication and assembly. Of course, such advances in the metrology will also have the advantage that they would minimize the handling of the mirrors. Finally, the metrology must be compatible not only with the fabrication tolerances but also with the spatial frequencies of the polishing techniques and tools used to achieve them.

The principal mechanical assembly and alignment criteria of the AXAF mirror assembly are: the ability to survive the mechanical, thermal, and acoustic loads during launch, and possibly reentry and relaunch; to maintain alignment during orbital operations; to X-ray test on the ground in the horizontal position and to permit meaningful prediction of the results of these tests to zero gravity; and to permit the assembly and alignment within specifications.

Three basic mirror support concepts have been examined by the study team which involves both NASA's Marshall Space Flight Center (MSFC) and the Smithsonian Astrophysical Observatory (SAO). These three concepts are illustrated in Figure 11. The cantilevered support concept is an extension of the HEAO-2 approach and lends itself to simple assembly and alignment procedures. However, it results in significant distortions of the optics under 1 g loading. On the other hand, the center support cylinder, which reduces the gravitational distortions, appears to complicate the assembly and alignment. The modified cantilever appears to combine the advantages of both approaches. All these designs apparently are viable, but considerations such as the preceding clearly indicate that a mere "scaling-up" of the HEAO-2 design must be treated with caution.

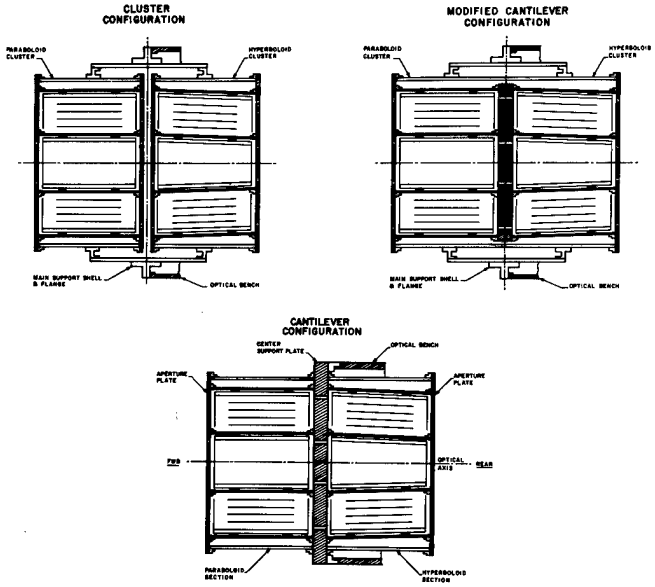


Figure 11. AXAF mirror support concepts.

Advanced technology studies

An advanced technology program for the AXAF X-ray optics is currently in progress at MSFC and SAO and also with contracts to Eastman Kodak, Itek and Perkin-Elmer. These studies have been undertaken to identify and understand the X-ray scattering properties of various candidate materials for the AXAF mirror elements and to study fabrication and metrology methods and techniques which may be applicable to the AXAF telescope.

Polished flat samples of a large number of materials, including fused silica and Zerodur (the leading candidate materials for the AXAF) have been fabricated, polished and coated and are currently being tested at the X-ray test facility at MSFC. A diagram of the test setup is shown in Figure 12. The second X-ray reflection was introduced only to render the emerging X-ray beam parallel to the axis of the vacuum chamber and thus amenable to detection at the location of the detector. The X-ray detector is a 1 inch diameter microchannel plate detector (the brassboard prototype of the HEAO-2 high-resolution imaging X-ray detector) and allows us to measure X-ray scatter at the sub-arc-second level. It is interesting to note that, in the absence of any scattering effects, the detected radiation pattern is dominated by single-slit Fresnel diffraction.

In addition to the X-ray tests, the flat samples are also undergoing a series of measurements using a wide variety of metrological techniques (Table 3) in an attempt to determine quantitative relationships between the X-ray performance and the surface topography. Optical metrology for the aspheric mirror surfaces is also under study and breadboarding.

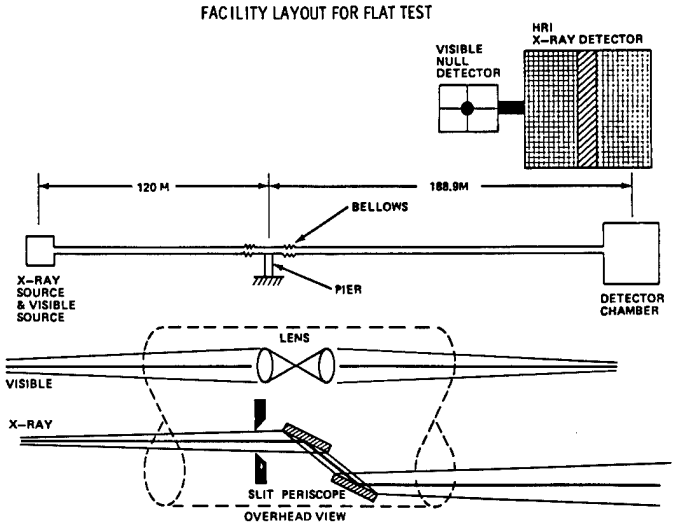


Figure 12. Schematic diagram of the test setup being used to measure X-ray scatter from various samples.

Table 3. Tests to be Performed on AXAF Optical Flats

Test	Purpose/Description
Twyman-Green Interferometry	Determines flatness
Nomarski Photography	Qualitative evaluation of surface topography
FECO	Quantitative determination of rms surface roughness and correlation length
Talystep	Contact method of determining surface topography
Total Integrated Scatter	Visible light method for determining rms surface roughness
Visible Light Scatterometer	Visible light method for determining rms surface roughness and autocorrelation length
Polarization Profilometer	Non-contact method of determining surface topography

Finally, as part of the technology program and the phase B activity (currently scheduled to start in fiscal 1981), each of the competing systems contractors will build a 40 cm diameter Wolter type I paraboloid/hyperboloid. The 40 cm diameter was chosen to allow grazing angles which would yield an energy response over the full AXAF range and, at the same time, allow the mirror to be tested in X-rays at MSFC without any modification to the test facility. These test mirrors will be fabricated with the current AXAF tolerance specifications and will serve to test techniques and demonstrate the methods proposed to meet the AXAF requirements.

Spacecraft considerations

The great majority of the spacecraft requirements for the AXAF are noncontroversial and well within the current state-of-the-art. These requirements and their feasibility using existing components have been discussed in detail in the MSFC/SAO Phase A reports and will not be repeated here. The area requiring the most advanced technology involves the aspect determination.

The typical X-ray observation with the AXAF will require more than 1000 seconds to form an image. During this time there may be significant motion of the spacecraft relative to the target and/or the motion of the X-ray detector relative to the X-ray telescope. Fortunately, X-ray detectors are single photon counting devices. Therefore, to reconstruct the X-ray image, one needs to know only the instantaneous position of the X-ray telescope axis and the relationship of this axis to the detector at the instant any particular event is detected, thus eliminating the requirement that these axes be absolutely stable. The required knowledge of the positions of these axes will be obtained by means of aspect star cameras to monitor the alignment of the telescope axis to the celestial sphere and a fiducial light system to monitor the alignment of the telescope with respect to the detectors. The optimal arrangement of this monitoring system would have the fiducial lights projected back through the telescope and onto the star camera.

In order to meet the scientific requirements of the AXAF the present goal is to achieve an aspect solution with an absolute accuracy of 1.0 second of arc and a relative accuracy of better than 0.5 second of arc. These requirements imply a fine aspect camera with a noise equivalent angle of less than 0.3 second of arc. A requirement of a 95 percent probability of finding enough stars for an aspect solution anywhere on the celestial sphere and the present magnitude limit of accurate star catalogs imply that the aspect camera must be sensitive to ninth magnitude stars and have a total field of view of $\sim$4 square degrees. The present view is that a star camera using a charge transfer device, either CCD or CID, should be built to meet these requirements.

Scientific instruments

NASA is planning to release an Announcement of Opportunity (AO) for scientific investigations, some of which would include the design and development of instruments. The AO is planned for release some time this year. It is expected that the number of instruments selected for study during the design definition phase will be greater than the number finally chosen for the initial flight of the AXAF. Subsequent AO's will be released for the refurbishment and/or development of new instruments based on a nominal schedule of on-orbit maintenance visits to the AXAF by the Space Shuttle every 30 months.

During the past 2 years a Science Working Group (members are listed in Table 4) has been involved in the AXAF study. The group has identified the following observational objectives which are required to meet the scientific goals of the AXAF: imaging, spectroscopy, polarization and timing. These observational goals together with the scientific requirements and performance goals of the AXAF will require:

(a) Single photon counting X-ray imaging detectors with high spatial resolution, high quantum efficiency, and wide fields of view.

(b) Spectrometers with both high ($E/\Delta E \geq 1000$) and low ($E/\Delta E \gtrsim 10$) energy resolution and high quantum efficiency and moderate time resolution.

(c) Polarimeters of the highest sensitivity to X-ray polarization.

Table 4. Members of the AXAF Science Working Group

Professor Riccardo Giacconi - Harvard University - Chairman
Dr. Martin C. Weisskopf - Marshall Space Flight Center - Vice-Chairman
Dr. Elihu Boldt - Goddard Space Flight Center
Professor Stuart Bowyer - University of California, Berkeley
Professor George Clark - Massachusetts Institute of Technology
Professor Arthur Davidson - Johns Hopkins University
Professor Gordon Garmire - California Institute of Technology
Professor William Kraushaar - University of Wisconsin
Professor Robert Novick - Columbia University
Dr. Albert Opp - NASA Headquarters - ex officio
Professor Minoru Oda - Tokyo University, Japan
Professor Kenneth Pounds - University of Leicester, United Kingdom
Dr. Seth Shulman - Naval Research Laboratory
Dr. Harvey Tananbaum - Harvard/Smithsonian Center for Astrophysics
Dr. Joachin Truemper, Max-Plank Institute, Germany
Professor Arthur Walker - Stanford University

No single X-ray instrument can meet all of these requirements. Tables 5 and 6 list possible candidate instruments and their expected performance. In addition, the AXAF may also have onboard, nonfocal plane out-of-the-optical-path instruments such as an all sky monitor and/or a collimated monitor counter.

Table 5. Imaging X-Ray Detectors

Detector	Size (Field of View)	Spatial Resolution	Quantum Efficiency	E/ΔE @1 keV	E/ΔE @6 keV
Charge Coupled Device	≥25 mm x 25 mm (≥8 x 8 arc min)	15-25 μm (0.3-0.5 arc sec)	high	5	30
Negative Electron Affinity Detector	≥25 mm diameter (≥8 arc min)	15 μm (0.3 arc sec)	high	5	1
Microchannel Plate	≥85 mm diameter (>30 arc min)	15 μm (0.3 arc sec)	low	None	
Imaging Proportional Counter	180 mm x 180 mm (1^o x 1^o)	<0.5 mm (<10 arc sec)	high	2	5
Gas Scintillating Imaging Proportional Counter	180 mm x 180 mm (1^o x 1^o)	<0.5 mm (<10 arc sec)	high	10	12

Table 6. X-Ray Spectrometers and Polarimeters

Instrument	Energy Range (keV)	Energy Resolution (E/ΔE)	Location
Filter Wheel	0.1 - 8	3	in front of detector
Transmission Grating	0.1 - 4	100 - 200	behind mirrors
Solid State Detector	0.4 - 8	15 - 50	in focal plane
Objective Crystal Spectrometer	0.5 - 8	500 - 10,000	in front of mirrors
Focal Plane Crystal Spectrometer	0.5 - 8	500 - 3,000	in focal plane
Focal Plane Concave Grating Spectrometer	0.1 - 1	500	in focal plane
Gas Scintillation Proportional Counter	0.1 - 8	10 - 15	in focal plane
Bragg Crystal Polarimeter	2.6, 5.2	N/A	in focal plane

Summary

Broadly stated, the scientific goals of the AXAF are to:

- Determine the nature of celestial entities ranging from normal stars to quasars.
- Understand the nature of the physical processes which take place in and between astronomical objects.
- Determine the history and understand the evolution of the universe by means of X-ray observations.

The AXAF will be used to probe and understand the bursters, pulsars, and supernova remnants. It will provide information about the structure, composition, and dynamics of our own galaxy and the physical processes involved. X-ray observations with AXAF will permit the study of the dynamics and evolution of clusters of galaxies and extend these observations to the very faintest sources. AXAF will have the capability to observe quasars to the very edges of the universe ($Z \simeq 10$). These represent only a few of the astrophysical problems which AXAF is designed to address.

That the scientific goals of the AXAF can be so far reaching is a direct consequence of the spectacular success of the Einstein Observatory, which proved to be a significant milestone in the development of the field of X-ray astronomy. The Einstein extended the scope of this field to the point where all known classes of astronomical objects have been observed in X-rays. The results of the Einstein observations have thus allowed X-ray astronomy to take its place as one of the major tools at our disposal in the study of the universe. Yet, despite its success, it is sobering to realize that the light collecting power of the Einstein telescope was a step backward to the actual capabilities of the early development of the field and that the angular resolution was severely limited. The Einstein was clearly necessary to prove a new technology. This has now been accomplished, and in the intervening years even newer technology has evolved to provide us with the assurance that the AXAF can be built and utilized to its fullest potential.

Acknowledgments

The continuing progress on the AXAF study is due to a joint team of scientists, engineers, and managers at the Marshall Space Flight Center, the Smithsonian Astrophysical Observatory, NASA Headquarters, and the members of the Science Working Group. Dr. Martin Zombeck (SAO) and Dr. Charles Wyman (MSFC) were particularly helpful in the preparation of this paper.

Bibliography

MSFC Phase A AXAF Study Report
AXAF Science Working Group Report (in draft form)
SAO-AXAF Reports 79-002, 79-005, 80-003

A large-aperture space telescope for infrared and submillimeter astronomy

James P. Murphy, Michael K. Kiya, Michael Werner
Ames Research Center, NASA, Moffett Field, California 94035
Paul N. Swanson, Thomas B. H. Kuiper, Paul D. Batelaan
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California 91103

Abstract

Large apertures improve the capability of astronomical telescopes in two ways: increased spatial resolution (linearly dependent on aperture size) and increased sensitivity (dependent on aperture size squared). For the purpose of a technology assessment, the Large-Aperture Telescope (LAT) for infrared and submillimeter astronomy is envisioned to be 10 to 30 m in diameter, operating in the 2-μm ≤ λ ≤ 1000-μm wavelength range. It would be carried to orbit with a single launch of the Space Transportation System and semi-automatically deployed as a free flyer with a nominal 10-yr mission duration. Periodic revisits at 2-yr intervals would allow servicing and instrument change. LAT must be placed above the Earth's atmosphere to avoid the absorption that occurs through much of the infrared and submillimeter, and to avoid turbulence which limits spatial resolution. Important technical considerations for LAT include: (1) telescope optical form, (2) primary mirror material, (3) figure control techniques, (4) deployment techniques, (5) pointing and stabilization, and (6) thermal control. This paper discusses the science objectives and rationale for LAT and describes different hardware techniques and concepts for its implementation.

Introduction

NASA is currently in the early stages of planning for a potential space-based Large-Aperture Telescope (LAT) for infrared and submillimeter astronomy. LAT is a multi-segmented deployable telescope with a collecting diameter of 10 to 30 m and a capability for housing multiple scientific instruments. A technology assessment study for such a system is presently underway with Lockheed Missiles and Space Corp. under the joint direction of NASA-Ames Research Center and the Jet Propulsion Laboratory. LAT would provide a significant advancement in sensitivity and resolution for astronomical observations in the infrared and submillimeter wavelengths. Figure 1 shows the potential of a LAT with a nominal 15-m collecting aperture versus several current or planned NASA telescopes. The Kuiper Airborne Observatory is a 0.9-m telescope, housed in a NASA-owned and operated C-141, which typically makes observations at 41000-ft altitude. The IRAS is the Infrared Astronomy Satellite, a 0.6-m aperture cryogenically cooled telescope which will survey the sky in the 8- to 120-μm spectral region. It is currently under development to be flown as a free flyer in the 1982 time frame.[1] The ST is the Space Telescope, a 2.4-m aperture visible-wavelength telescope, which is also currently under development and scheduled to be flown as a free flyer in the 1983 time frame. The SIRTF is the Shuttle Infrared Telescope Facility, a 1-m aperture, cryogenically cooled telescope designed for operation from 2 to 1000 μm. It is currently under study and is scheduled to be flown in the Space Shuttle sortie mode beginning in 1986.[2] In comparison to these systems, LAT's greater aperture provides much better resolution in all cases and, except for SIRTF, much greater sensitivity. It will therefore open up many new scientific areas for study. For example, at 100 μm the LAT angular resolution of ~1" is more than an order of magnitude gain over the best currently available in this spectral range and is, in fact, comparable with the angular resolution used for present-day ground-based optical and near-infrared observations. The LAT may also provide the first opportunity for high-resolution spectroscopic and positional follow-up of many of the exciting discoveries of the IRAS survey.

In the next section, the advantages of space for infrared and submillimeter astronomical observations are outlined. This is followed by summaries of current thinking concerning scientific objectives, potential scientific instruments, baseline requirements, and technical considerations for the LAT.

Advantages of space for infrared and submillimeter astronomy

Atmospheric absorption

From about 30 μm to 700 μm, the atmosphere is almost totally opaque from even the highest mountain-top observatories,[3] and outside of this range there are additional bands of

forbiddingly high opacity (Figure 2). The 30-μm to 700-μm band encompasses the peaks of blackbody radiation curves for objects with temperatures in the range of approximately 5 to 150 K as well as many atomic fine-structure lines and molecular rotational transitions. Consequently, it is the wavelength range best suited for the study of an enormous variety of astrophysical phenomena, but telescopes above the atmosphere must be used to exploit it. Airborne platforms provide a substantial (but not total) reduction in atmospheric absorption, but the telescopes which can be lifted to such altitudes are small, on the order of 1-m diameter. By way of comparison, the adjacent millimeter wave band has several telescopes in the 10- to 15-m class with 25-m telescopes planned for the 1980's, while the optical and near-IR band has 4- to 5-m class telescopes, with a 10- to 15-m class instrument envisioned for the next decade. A 10- to 20-m space telescope will provide dramatically improved capabilities for infrared and submillimeter observations and also bridge the gap between millimeter and optical wavelengths, thereby maintaining our essential ability to study new astrophysical phenomena across the entire spectrum.

Atmospheric emission

Ground-based observations through the millimeter and submillimeter windows are hampered by the variability in the emission from the Earth's atmosphere on relatively short time scales (ms to min). The emission also varies spatially over small angles, so that differential measurements of two adjacent positions on the sky cannot entirely overcome this limitation. The seriousness of these limitations depends on the measurement desired. As one example, ground-based measurements subject to these limitations have set upper limits of 0.01% on the fluctuations in the cosmic background, but further progress on this crucial scientific problem may require a space telescope free of atmsopheric emission effects.

Atmospheric seeing

A telescope outside the atmosphere will not be affected by the atmospheric fluctuations on scales less than or equal to the telescope aperture which limit the resolving power of large optical and infrared telescopes on Earth. At a wavelength of 30 μm, a 15-m diameter LAT will achieve a resolution of ~0.5", comparable to the very best images of ground-based optical telescopes.

Structural limitations

There is a limit to the size of structures which can be erected on the Earth's surface, due to the strength of available materials. On the other hand, there is no physical reason why a 1000-m optical telescope could not be built in space. At issue is not rigidity, since the reflectors can be corrected if they deviate from the ideal shape, but simply strength. Thus, the space environment will not only free us from the limitations imposed by the atmosphere, but also permit the use of types of telescopes which simply cannot be constructed on Earth.

Science objectives

The large collecting area, high angular resolution, and access to a largely unexplored spectral region with total freedom from atmospheric interference would give the LAT unique capabilities, so it could carry out classes of observation not otherwise possible. Many of the scientific problems it would study cannot yet be defined. In terms of our current knowledge of astrophysics, however, it is possible to describe many areas where it would have a major scientific impact. Although a detailed scientific justification will not be presented here, it is clear that the capability which this telescope offers of making observations of high spectral resolution with high spatial resolution throughout the far-infrared and submillimeter band will make it a very powerful tool for the study of star formation both in our own galaxy and in nearby galaxies.

Star formation

This point is illustrated in Figure 3, which depicts the region W3, which is one of the best-studied regions of current star formation in our galaxy. The left panel shows contours of far-infrared ($\lambda \sim 100$ μm) continuum emission, measured with 5' resolution, superimposed on a photograph of this region. The far-infrared luminosity is $\sim 10^6$ solar luminosities, and the continuum emission appears to be thermal radiation from interstellar dust particles. The rectangle at the center of the far-infrared map contains no notable optical features, but observations at near-infrared ($\lambda \sim 20$ μm) and radio continuum ($\lambda \sim 6$ cm) wavelengths (right panel) have shown this small region to contain a cluster of interesting and unusual objects which appear to be young and forming stars in a variety of early evolutionary stages. These young stars are embedded within, and are the heat sources for, the dust clouds which radiate the large far-infrared luminosity of the W3 region.

The near-infrared and radio observations of the W3 cluster shown in Figure 3 were made with an angular resolution of a few arc sec which permits the objects in the cluster to be separated spatially to be studied in detail. Also shown in the right panel is the distribution of far-infrared emission over this central region, as determined with the best spatial resolution currently attainable, 20" from NASA's Kuiper Airborne Observatory. At this resolution, the region shows little structure, but with the 1" to 2" resolution attainable with the LAT in the far-infrared, it will be possible for the first time to study the individual objects in very young clusters like this. These high angular resolution observations will include broadband measurements to determine the luminosities of the separate sources and also high spectral resolution observations to study the composition, temperature, density, and motion of the gas clouds associated with these sources. The far-infrared and submillimeter spectral region is rich in spectral lines of molecular (e.g., CO, H_2O), neutral atomic (O, C), and ionized (O^{++}, N^{++}) species which are produced under varying conditions and thus serve as probes of a variety of astrophysical environments; moreover, observations in this spectral band are capable of penetrating through the dust clouds to sample the gas most intimately associated with the young stars. Spectral line observations of the sources in the W3 cluster can therefore investigate such questions as the degree of ionization produced by the embedded sources, the importance of shock waves in promoting or inhibiting star formation, and the presence or absence of infalling or outflowing circumstellar matter. These observations will lead to improved understanding of the properties of the individual sources and their interaction with each other and with the environment from which they have formed. This information in turn will be used to contrain and refine theories of star formation in the galaxy.

Galactic studies

The spatial scale of the region shown in Figure 3 is ~30 light yr, characteristic of luminous regions of star formation in the galaxy. A very exciting prospect for the LAT is that it will be capable of both detecting and resolving regions such as this in nearby galaxies, as shown in Figure 4, and will thus permit a systematic study of star formation in other galaxies. This can be advantageous because these galaxies are seen from a favorable external vantage point, so that their large-scale characteristics can be discerned more readily than for our galaxy. Figure 4 illustrates that the high angular resolution of the LAT, comparable to what is attained with large, ground-based optical telescopes, will resolve a nearby galaxy into many spatial resolution elements. Studies of molecular emission on this fine scale will reveal the abundance and properties of clouds of molecular gas from which stars are thought to form, their motion and distribution relative to spiral arms and other features thought to be related to star formation, and the presence of highly excited gas clouds which might contain groups of recently formed stars. Far-infrared broadband observations on the same spatial scale, also carried out from the LAT, would complement the spectroscopic measurements by determining the luminosities of the young and forming stars. The LAT will permit studies of this type of a large sample of galaxies. Thus, the pattern and extent of star formation can be correlated with other galactic properties, such as mass, rotation velocity, and morphological type, to indicate how these factors interact to influence the structure and evolution of galaxies.

Galactic and extragalactic star formation is, of course, not the only scientific area where the impact of the LDR will be considerable. Briefly, several other important problems it can attack are the following:

Small scale fluctuations of the microwave background

Clumpiness in the distribution of the "big bang" fireball radiation is one of the few means of exploring phenomena such as turbulence and density perturbations in the primordial universe which are of fundamental importance in understanding galaxy formation. Present ground-based observations at millimeter and centimeter wavelengths show that the distribution is smooth to a level of at least $\Delta T/T \simeq 10^{-4}$ on spatial scales from 1' to 100'. The LAT, operating at 1 mm with $\Delta\lambda/\lambda = 0.4$ and a 1' field-of-view, can set a 10 σ limit on $\Delta T/T$ of 10^{-6} in 10^3 sec. Both the spatial scale and the sensitivity of such a measurement are in the range where many theories predict that fluctuations should be detectable. These types of measurement can also be used to probe the intracluster medium in X-ray emitting galaxy clusters by looking for heating of the background radiation due to scattering by relativistic electrons. The telescope will be free of atmospheric effects which may hamper further ground-based progress in this field; additionally, its access to the entire spectrum and the possibility of making measurements on a range of angular scales should make it possible to separate the effects of interest from purely local ones arising within our galaxy.

Extragalactic radio sources and galactic nuclei

While it is generally accepted that quasars, the nuclei of radio galaxies, BL Lacerta objects, Seyfert galaxies, etc., emit large amounts of radio radiation by the synchrotron mechanism, the details of this mechanism, the physical conditions which give rise to it,

and the temporal variations and evolution of such objects are not understood. Some of these objects have spectra whose flux continues to rise at the shortest observed wavelengths (1 to 3 mm). Many are powerful emitters in the infrared. Many of these objects show variability in their radio emission. Similarly, variability is seen at infrared and visual wavelengths. In order to improve our understanding of these objects, the radio phenomena and infrared phenomena must be linked. This requires observations at far-infrared and submillimeter wavelengths where, at present, only a handful of such objects have been detected. In a related problem, the high spatial resolution of the telescope can be brought to bear on the nuclei of galaxies which show powerful far-infrared emission. For example, the central region of a nearby galaxy like M82, which has a luminosity of 4×10^{10} solar luminosities peaking at 100 μm, can be mapped on a spatial scale of 75 light yr. This would show whether the emitting material is clumped or smoothly distributed. These two results have differing implications concerning the source of the luminosity and the triggering of the high luminosity phase. If the high angular resolution mapping uncovers regions which appear to be sites of enhanced star formation, high angular resolution spectroscopy of the associated atomic and molecular gas would yield crucial diagnostic information about velocities, composition, and excitation conditions.

Comets and planets

The LAT will be a powerful tool for the investigation of objects in our solar system. Comets, for example, are believed to consist of ices and dust left over from the time the solar system was formed. The ices are believed to include relatively complex molecules which sublime when the comet approaches the sun. Within recent years, some such molecules have been detected by their microwave and millimeter wave transitions. The signals received are very weak, but indicate that the molecules are highly excited. It is most probable that transitions in the submillimeter range will be much stronger, thus permitting useful quantitative studies to be conducted. If we can come to understand the molecular composition of comets, we may then be able to relate this to the composition of protostellar clouds. A good understanding of the gaseous composition of the protosolar nebula would, in turn, serve as a stepping-stone to an understanding of the evolution and current composition of planetary atmospheres. Microwave transitions are also useful in the study of planetary atmospheres, particularly in determining the vertical variation of temperature and composition, and the high angular resolution of the LAT would resolve the disk of the giant planets into many resolution elements to permit zonal variations in the atmosphere to be studied.

Science instruments

A number of instruments would be used to cover the broad spectral range of the LAT and to encompass the large variety of scientific problems, and it is anticipated that several instruments would share the focal plane and the observing time. Except at the shortest wavelengths, the thermal emission from the telescope will prove the main source of background, and background-limited observations will be possible with broad spectral bands ($\Delta\lambda/\lambda \gtrsim 0.1$) for detector NEPs below 10^{-15} W/$\sqrt{\text{Hz}}$ (see Figure 5). This performance can be achieved currently with ^{3}He bolometers at submillimeter wavelengths and is exceeded by 1- to 2-orders of magnitude by currently available photoconductors at wavelengths $\lambda \lesssim 250$ μm. It is therefore clear that optimum instrumentation for use of the LAT with low-to-moderate spectral resolution is readily achievable with current technology. Monolithic arrays of infrared photoconductors would be ideally suited for imaging at the high spatial resolution of the telescope, as well as for use with dispersive spectrometers. Such arrays are just starting to become available to infrared astronomers; it seems likely that this technology will progress rapidly. For the highest spectral resolution, heterodyne (coherent) receivers are the logical choice. Recent progress has pushed these techniques up to 500 GHz (600 μm) with adequate performance, and these systems are routinely used around 1 mm for astronomical observations at a number of telescopes. Progress in this area, including extension to shorter wavelengths, is being made very rapidly. To summarize, radiometric and spectroscopic instruments of the type which would be used on the LAT are in continual use and/or under rapid development for such related NASA programs as IRAS, SIRTF, and the Kuiper Airborne Observatory, as well as for ground-based observing in the near-infrared and the submillimeter. These and related efforts will lead to the development of a suitable complement of scientific instruments in parallel with the development of the telescope itself.

Systems requirements summary

The LAT system requirements are summarized in Table 1. It should be emphasized that these parameters represent a starting point for systems studies and will be refined as LAT technology and design evolves. The baseline for launch and deployment is a single flight of the Space Transportation System (STS) with a semi-automatic deployment, although automatic deployment is highly desired. The mission duration would be in the order of 10 yr with periodic revisits at 2-yr intervals for maintenance, replacement of instrument cryogens, and exchange of instrument complements. The telescope collecting apertures being considered are from at least 10 m, up to 30 m diameter. The operating wavelengths of the telescope

will be from the near-infrared (2 μm) to submillimeter (1000 μm). The optics are to be diffraction-limited at approximately 30 μm, with 2 to 10 μm highly desired. Diffraction limit versus degree of active control versus aperture size will be important tradeoff areas in arriving at the final LAT systems requirements. The telescope must be capable of slewing and stabilizing on a celestial source from another over 180° in 10 min with 5 min highly desirable. To reduce the background radiation, spatial chopping over the telescope field-of-view at rates up to 10 Hz is required. This chopping could be accomplished with the secondary mirror or with additional optical elements in the focal plane of the telescope. Non-rejected stray light radiation from any source ≥60° off-axis must be less than the combination of telescope self-emission and natural (i.e., zodiacal) backgrounds. Telescope cooling requirements are presently under study, and it is anticipated that passive cooling to 150 to 200 K will be desirable.

Technical considerations

The LAT will be a parabolic reflector of 10 m or more in diameter. It must have an rms surface tolerance of less than 2 μm to be usable at a 30 μm wavelength. It must be capable of being deployed from the Space Shuttle which has a cargo bay diameter of 4.5 m. Figure 6 shows a representative 10 m concept for LAT being deployed from the shuttle.

This section discusses some of the technical considerations involved in defining a LAT system and indicates those areas where further analysis and tradeoff studies appear to be needed.

Optical configuration

The only telescopes which are completely wavelength independent are reflectors, and the choice for the LAT narrows to the type of reflecting telescope to be used. Several reflecting telescope configurations have been considered, including spherical and parabolic primaries with both Cassegrain and prime focus, Ritchey-Chretien systems, and a tilted aplanat. For the current LAT requirement of a small field of view, the parabolic primary has the smallest aberration of any of the reflectors considered and provides diffraction-limited operation across the specified field of view. A Cassegrain configuration appears to be the most desirable, since the detectors can be near the vertex of the parabola. This provides easy access to spacecraft systems, such as cryogenics and thermal control, and makes it easier to mechanically support a large instrument package.

Reflector design and fabrication

Figure 7 is a plot of telescope diameter versus wavelength. The "shaded" area labeled "new technology" is the regime of the LAT. The points 1 to 11 on the right of the figure represent the minimum usable wavelength for the best ground-based lightweight reflector technology. The solid diagonal line shows that, generally, the diameter is about 10^5 times the rms surface tolerance. If the criterion of a minimum operating wavelength of 16 times the surface tolerance is adopted, then one can say that the maximum reflector diameter is about 6000 wavelengths, again independent of wavelength. At the left of Figure 7 are two points, 12 and 13. These represent a technology called "monoliths" which are typical of traditional optical mirror construction. Monoliths have an areal density more than an order of magnitude greater than the "lightweight" radio telescopes.

Figure 7 shows that the LAT parameters differ from those of previous telescopes, and that new technologies will be required in its fabrications. An area of particular interest is the areal density of the reflector, because of its impact on the weight of the entire system.

Figure 8 addresses the weight problem in terms of the Shuttle capability. The unassisted Shuttle can place 32000 lb into a 900 km equatorial orbit. Assuming 10000 lb for the spacecraft package (comparable to the Space Telescope), the remaining 22000 lb are available for the reflector plus deployment mechanisms and associated hardware. Assuming a 10% weight overhead for deployment, a 12-m diameter reflector must have an areal density of less than 170 lb/m^2, and a 17-m reflector must be less than 75 lb/m^2. The areal densities of several types of reflector technology are also shown.

Figure 9 shows four different concepts for the reflector. The lightweight monolith is not really lightweight, the Space Telescope mirror being 400 lb/m^2. However, advanced frit-bonding techniques may reduce the areal density to less than 200 lb/m^2 sometime in the future. The maximum diameter of a monolith is limited by the Shuttle bay diameter to 4.2 m. However, this technology may be useful for the truss and segment approach to be discussed below.

The membrane antenna, pioneered by both the General Research Corporation and MIT, is an ingenious approach that uses a non-precision, conducting backup structure. A flexible

conducting membrane is stretched over the front of this structure and pulled in by electrostatic forces. By dividing the conducting surfaces into many smaller conducting islands and providing a charging path to each island, the reflector figure can be continuously controlled. Although extremely light and relatively easy to deploy, this technology is not yet mature and is not presently being considered for the LAT.

The shell reflector, without a backup structure of the type developed by TRW, has an areal density of less than 25 lb/m^2. Deployment is straightforward, and the design is mature; however, presently obtained surface accuracies are far short of the LAT requirement, and the addition of active surface control may present difficulties.

The remaining reflector in Figure 9, the truss and surface segment, appears the most viable for the LAT. A truss backup structure of high rigidity supports a number of hexagonal panels at three points per panel. This design is readily adaptable to active control by replacing the panel supports by linear actuators. The LAT would use panels or segments of approximately 4-m diameter. Seven close-packed hexagons would form a 10-m diameter reflector, and 19 hexagons a 17-m reflector.

The construction of the individual reflector segments depends on the minimum operating wavelength, and this is one of the major tradeoff areas to be studied. To operate at a wavelength shorter than about 50 μm, optical mirror techniques may have to be used. An adaptation of the Space Telescope "monolith" approach would provide the surface accuracy in a completely passive segment but at a very high weight. This construction technique is to sandwich a welded-glass egg-crate structure between a planar faceplate and backplate. The mirror faceplate is fused to the egg-crate structure by heating the faceplate to a temperature high enough to soften and fuse it to the egg-crate, but too low to destroy the mirror structure itself. The mirror is inverted, the backplate is similarly fused, and the entire mirror simultaneously slumped to a best-fit sphere. The faceplate may then have to be polished. This type of construction requires a relatively thick faceplate, backplate, and egg-crate structure, not for figure preservation or strength, but to endure the thermal environment of the manufacturing process.

The development of a low temperature (900° C) glass-like bonding agent (frit bonding) for Ultra Low Expansion (ULE) glass relieved the high-temperature, high-risk steps in the manufacturing process. Much thinner structural components are possible (frontplates of 1/4 to 1.3 in., egg-crate thickness of 1/20 in.). The frit bonding technique makes possible lighter mirrors than the technology of glass welding. Frit technology produces mirrors which are not only lightweight but also stronger than the conventional technology because of the better bond of core to frontplate. The frit mirror is, thus, better able to withstand launch loads.

For operating wavelengths longer than ~100 μm, honeycomb segments, such as used by Leighton in construction of his 10-m reflector which has a surface tolerance of 10 μm, may be used. Skins of graphite epoxy can be bonded to both sides of a light honeycomb core. This type of structure can have an areal density of less than 30 lb/m^2 for passive segments of up to 4-m diameter. The problem of the surface finish on the graphite epoxy has yet to be addressed, although machining into a gel coat or electroplating a metallic surface are possibilities.

Additional possibilities for the reflector construction are under review because of rapid advances in lightweight materials technology.

Adaptive optics

It is almost a certainty that a completely passive reflector cannot be deployed to achieve and maintain a few micron surface accuracy over a long time and changing thermal conditions. Adaptive optics is the term applied to a system which continuously determines the relative position of all the optical elements, or determines the phase-path errors over the aperture. This system also either adjusts the relative position of the optical elements or adjusts the phase path to ensure that the optics remain diffraction limited. Figure 10 is a schematic representation of a system which measures the figure of the primary reflector by a laser rangefinder or, alternately, by relative-position sensors on each segment edge. Three actuators on each segment adjust the segments in two axes of tilt and in piston (movement normal to the surface). Reaction forces of the actuators act on the backup structure. Laser rangefinders and edge sensors are available which give unambiguous measurements to a fraction of a micron. If, because of thermal distortion, material creep, or any other reason, the required surface tolerance cannot be maintained by adjusting the relative positions of individual rigid segments, it may be necessary to provide another level of active control on each segment. Figure 11 is a schematic cross section of an illustrative reflector configuration showing the thin reflecting facesheet, an array of figure actuators to provide elastic control of the facesheet, a facesheet stiffener, and the segment actuators supporting each segment from the truss backup structure. The truss

structure itself may require actuators, not to control its shape, but for active damping to control normal modes excited by slewing or by vibrations from motors or pumps. Some possible materials are also indicated in Figure 11. For operation at very short wavelengths, or for a very lightweight reflector, both levels of actuators may be required. For a relaxed wavelength requirement, the surface figure actuators could be eliminated. For operation in the submillimeter region, the surface might in effect be the honeycomb stiffener itself; and above some wavelength, the segment actuators could be eliminated for a purely passive reflector. The degree of active control required for various reflector designs and operating wavelengths will be the subject of detailed study.

Other technical considerations

Deployment of the reflector from the Shuttle bay is a large but fundamentally solvable problem. Collapsible truss structures have been designed, and the surface segments can be individually attached to the backup by the remote manipulator. It may even be possible to implement a total self-deploying scheme. Deployment considerations will undoubtedly influence the final reflector design.

Because of the high angular resolution that will be obtained at the shortest operating wavelength, fractional arc-sec pointing is required. However, star trackers and servo loops with this accuracy already exist, and, the pointing requirement is no more severe than that of the Space Telescope.

Thermal control of the reflector is important for two reasons. First, the coefficient of thermal expansion and its uniformity across the surface will influence the degree of active control required. Temperature changes and gradients must be kept to a minimum in any case. Second, a reduced reflector temperature will reduce the background radiation, increasing the IR sensitivity. Methods of thermal control include thermal blanketing, sunshades, proper orbit selection, and selective coatings.

Summary

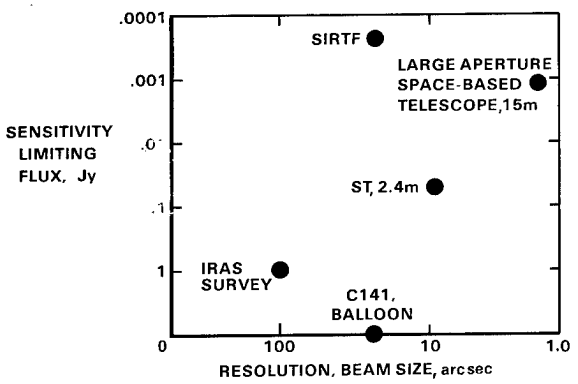
A space-based large aperture telescope offers an enormous increase in observational capabilities over any existing or currently planned telescope operating in the infrared and submillimeter wavelengths. Such a telescope will greatly enhance our understanding of star formation, galactic formation and evolution, extragalactic sources, comets, and planets. The advent of the Space Shuttle era combined with major advancements in lightweight optics and structures makes a space-based telescope many times the size of the 5-m Palomar telescope a possibility in the late 1980's. NASA is in the preliminary phases of defining such a system. As this definition becomes more developed, key technical areas to be considered are optical configuration, reflector design, adaptive optics, deployment, pointing, and thermal control. On the basis of our studies to date and progress on related programs, we feel that it is realistic to contemplate initiating the construction of this exciting and challenging telescope system before the end of the 1980's.

References

1. Aumann, H., and Walker, R., Opt. Eng., Vol. 16, No. 6, pp. 537-543. 1977.
2. Witteborn, F. C., Young, L. S., and Miller, J. H., SPIE, V. 183, p. 24. 1979.
3. Traub, W., and Stier, M., App. Opt., Vol. 15, No. 2, pp. 364-377. 1976.
4. Harvey, P. M., P.A.S.P., 91, p. 143. 1979.
5. Furniss, I., Jennings, R. E., and Moorwood, A. F. M., Astrophys. J., Vol. 202, p. 400. 1975.
6. Wynn-Williams, C. G., Becklin, E. E., and Neugebauer, G., Monthly Notices Roy. Astron. Soc., Vol. 160, p. 1. 1972.
7. Harper, D. A., personal communication. 1979.
8. Infrared Detectors for Low Background Astronomy, NASA TM 78598.

Table 1. LAT Requirements Summary

Launch/deployment:	Single STS launch to Low Earth Orbit (LEO) semi-automatic deployment (man assisted)
Mission duration:	10 yr; periodic revisits at 2-yr intervals
Time frame:	1987 technology; 1993 launch
Primary collecting aperture, D:	10 m $\leq$ D $\leq$ 30 m
Operating wavelength:	1000 μm > λ $\geq$ 2 μm
Diffraction limited at	$\leq$ 30μm
Field of view:	The greater of 30 sec or 10 × airy disk diameter
Pointing stability:	$\leq 0.1\ \lambda_{min}/D$ for >20 min interval
Retargeting slew:	180° in 10 min (5 min desired)
Spatial chopping:	Chopping frequency, 10 Hz; throw = F.O.V.
Stray light rejection:	Stray light from Sun, Earth, and Moon >60° off axis is less than effect of telescope emission (ε = 0.1) and zodiacal light on the detectors.
Primary mirror temperature:	~150-200 K;
Temperature uniformity:	Scan noise of spatial chopping shall not degrade achievable instrument sensitivity
Total telescope emissivity:	ε $\leq$ 0.1



COMPARISON WITH OTHER FACILITIES
SENSITIVITY VS ANGULAR RESOLUTION
100μm, Δλ/λ~0.5
S/N = 10

Figure 1. A comparison of the sensitivity and angular resolution at λ = 100 μm of a 15-m LAT with those of other existing and planned NASA facilities. The sensitivity estimates are based on current experience for 1 hr of integration for the C-141,[4] on projected performance for the IRAS,[1] and assuming background-limited performance, 30-min integration time, and an instrumental efficiency of ~0.1 for SIRTF, ST, and the large aperture space telescope.

ATMOSPHERIC TRANSMISSION

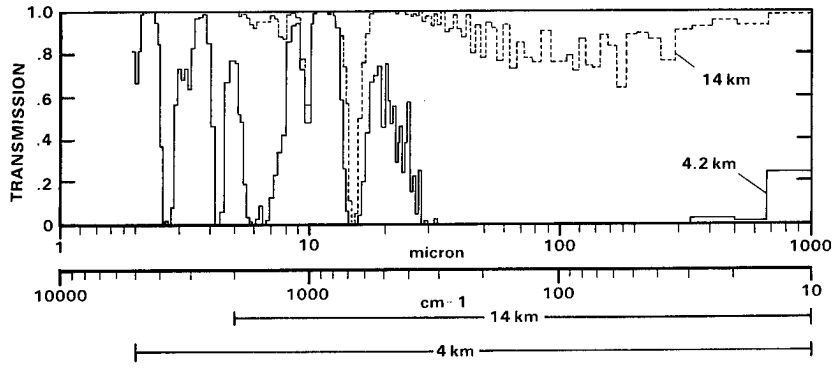


Figure 2. Atmospheric transmission from 1 μm to 1 mm at altitudes of 4.2 km, representative of a mountain top observatory, and 14 km, representative of an airborne observatory.[3] The bandpass varies from 5 cm^{-1} to 100 cm^{-1} to keep the resolution on the figure roughly to constant. Most of the absorption over this band is due to water vapor; a column density of 2.4 precipitable mm is assumed for 4.2 km altitude, and 4.5 precipitable μm for 14 km.

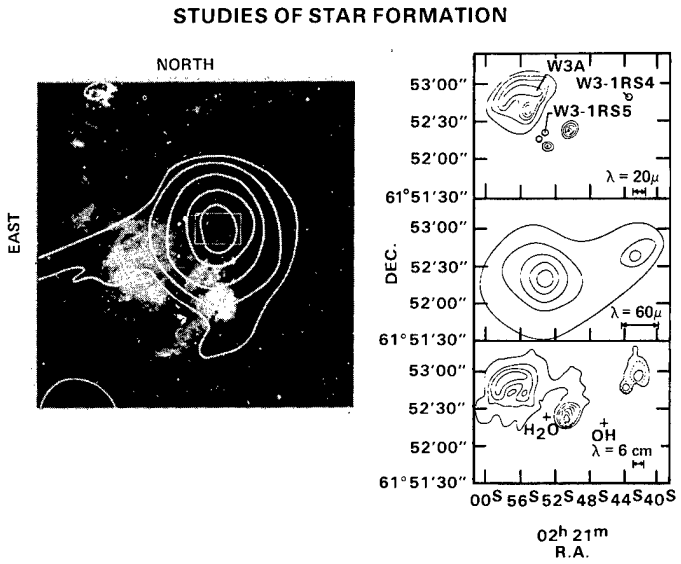


Figure 3. Star formation in the W3 region. The left panel shows the contours of far-infrared emission, as measured with 5' resolution[5] superimposed on an optical photograph of the W3 region. The box at the center of the 100-μm contours is shown on the right side of the figure as observed with ~5" resolution at 20 μm and 6 cm[6] and with 20" resolution at ~60 μm.[7] The LAT would make possible far-infrared measurements with beam sizes of a few arc sec.

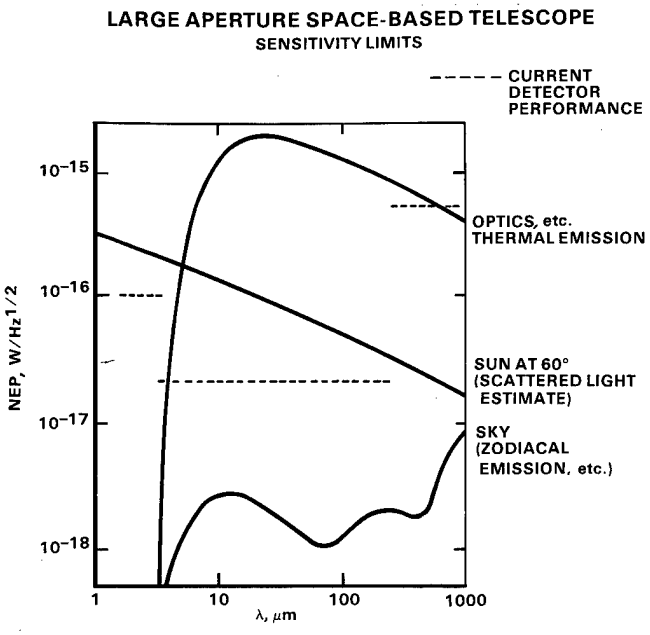


Figure 5. The sensitivity of currently available infrared detectors (cf. "Infrared Detectors for Low Background Astronomy" NASA TM 78598) is compared with the sensitivity limits set for the LAT by thermal emission and scattered light.

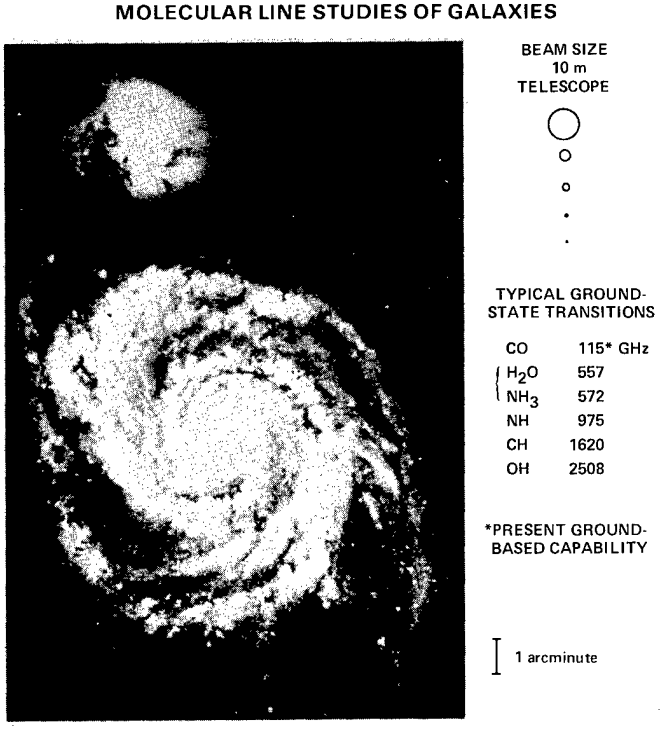


Figure 4. Molecular line studies of galaxies. The angular resolution of a 10-m LAT at the frequencies of a number of important molecular transitions is shown at scale relative to a photograph of the nearby galaxy M51.

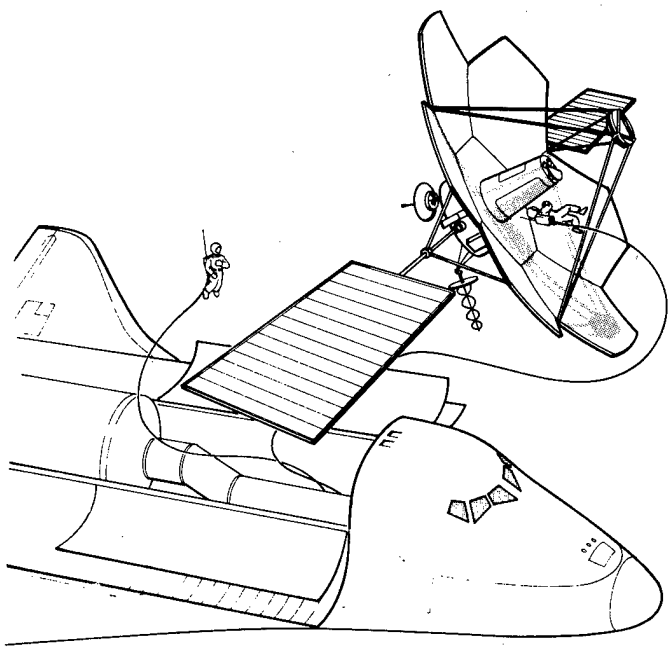


Figure 6. Artists conception of a 10-m LAT.

DIAMETER VS. OPERATING WAVELENGTH COMPARISON FOR THREE CLASSES OF TELESCOPES

1. U. TEXAS
2 OVRO 10 m
2A OVRO, ULTIMATE DESIGN GOAL
3 NRAO 11 m
3A NRAO 11 m, DOME CLOSED
4 CRIMEA
5 NRAO VLA
6 OVRO 40 m
7 BONN 100 m
8 NASA DSM 64 m
9 NRAO 90 m
10 PARKES 64 m
11 NRAO 25 m (PROPOSED)
12 HALE 200 in.
13 SPACE TELESCOPE

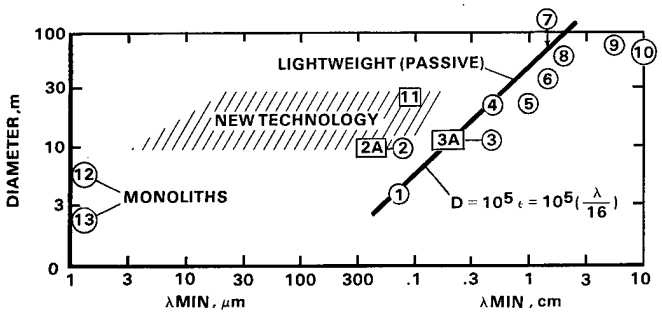


Figure 7. The minimum operating wavelength versus telescope diameter is shown for current optical telescope (monoliths) and radio telescopes (lightweight). The design goal for the LAT lies within the area designated "New Technology."

MIRROR TECHNOLOGY VERSUS SHUTTLE CAPABILITY

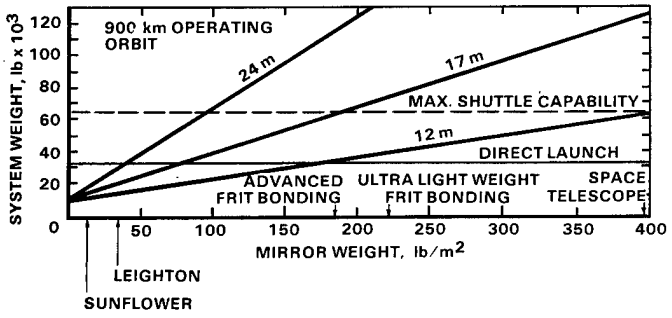
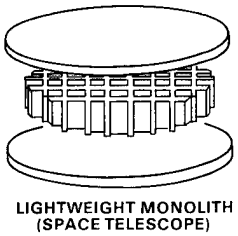
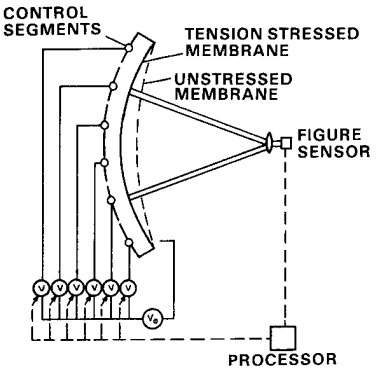


Figure 8. This figure compares the areal density of a reflector with the diameter of a telescope which can be placed in orbit by a single Shuttle launch under the assumption described in the text. The areal densities for several types of reflector technology are shown along the x-axis.

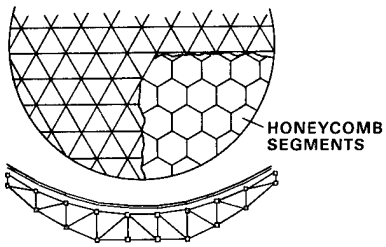
REPRESENTATIVE CONCEPTS



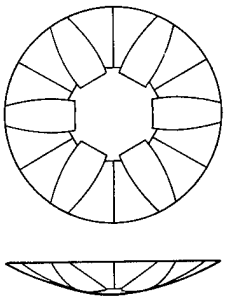
LIGHTWEIGHT MONOLITH (SPACE TELESCOPE)



MEMBRANE (GRC)



TRUSS AND SEGMENTS (LEIGHTON-CIT)



SHELL STRUCTURE (TRW SUNFLOWER)

Figure 9. Four representative concepts for the fabrication of a large aperture space telescope.

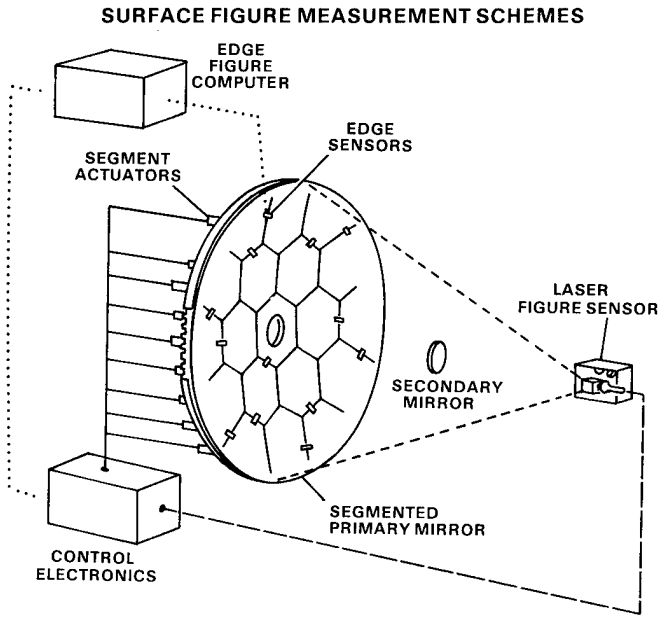


Figure 10. A schematic representation of edge- and surface-figure sensors for controlling the position of the panel in a segmented reflector.

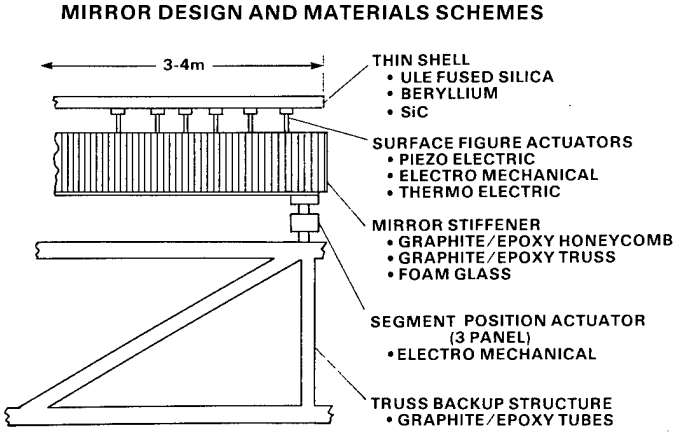


Figure 11. A schematic cross section through the surface of a hypothetical large-aperture space telescope, showing representative construction techniques, potential materials, and levels of control.

A very large space telescope for optical/UV astronomy

Max E. Nein, John W. Warner
Marshall Space Flight Center
Marshall Space Flight Center, Alabama 35812

Abstract

A concept is presented for a Ritchey-Chretien optical-UV telescope to be deployed in orbit using a preassembled segmented primary mirror of 8 meter aperture, carried into orbit inside the modified interstage of the Space Shuttle External Tank. Shuttle revisits allow major assembly tasks to proceed such as the installation of the light shield, fine alignment of the system and attachment of scientific instrument modules to the instrument adapter. Prime technology requirements for the VLST are assessed in this paper.

Introduction

The late 1980's and the early 1990's will be the era of the Space Telescope, our first national observatory in space. This long lived free flyer will expand significantly our knowledge in planetary, stellar, galactic and extragalactic astronomy. It will address the evolution of the universe as well as its current state, that is, the principles of order and the extent to which they apply to the matter in it. Space Telescope will provide a clear view of astronomical objects without the effects of absorption and "seeing" caused by the atmosphere. Its capacity for producing 0.1 arcsecond images, for imaging point sources to 27th visual magnitude and astrometry to 0.002 arcseconds accuracy will provide the best observing capability available for over a decade and is expected to meet the needs of astronomy until the mid-1990's.

However, the very nature of the scientific enterprise is to strive to reach the limits of technology and thus to extend knowledge to a maximum. It is therefore possible to foresee the day when higher resolution and larger collecting area will be needed to address the questions asked but not answered by the Space Telescope.

Science objectives beyond Space Telescope

Of the many such areas of astronomy which will require demonstrably better capability than current technology can provide, three will surely be mentioned: a further extension of the distance scale, a study of the more detailed structure of quasars and active galaxies, and a search for planetary systems around nearby stars.

An accurate determination of the size and scale of the universe through the determination of the local and global values of the Hubble constant H_o and the determination of the nature and fate of our universe through measurement of the deceleration parameter q_o is one of the major goals of the Space Telescope. Yet ST will not be able to study variable stars (primary distance indicators) to the Virgo cluster of galaxies and beyond. This cluster is a Rosetta Stone for extragalactic studies, since it is distant enough to fall within the Hubble flow regime of the universe but close enough to be known to possess a wide variety of galaxies and stars. A major goal of a successor to ST would be to determine the population, brightnesses, abundance and distances of the constituents of the Virgo Cluster in order to establish further a basis for studying the evolution and structure of the universe. Such a study requires several times higher resolution as well as a larger photon collecting area than ST can provide.

Current ground stationed very long baseline radio astronomy can study structures of radio sources with milliarcsecond resolution. Although ST can study the same sources in the optical/UV region the resolution will be fifty to one hundred times less because of the 2.4 m aperture size. While ST is expected to provide major new understanding regarding the nature and extent of energetic activity in galactic nuclei and its relationship to the phenomenon of quasars, direct comparison of the characteristics of active objects at comparable resolutions in such diverse areas of the spectrum as the radio, visible and the ultraviolet will not be possible. The scientific requirements for study of these objects beyond ST will again mandate higher resolution and more photons.

One of the goals of the Space Telescope is to measure the long term motions of the nearest stars over the lifetime of the mission in order to detect any low mass companions that might exist. Such a study may use both the direct astrometric capability of the fine guidance sensor on ST as well as the limiting resolving power of the two major cameras. The number of sources capable of being detected if they had companions is less than one hundred, thus providing a low probability of success. Several factors more in resolution would significantly increase the probability of detecting low mass companions, which is basic to the study of the possibilities of life existing in our galaxy.

Options for higher resolution and larger collecting area

Three possible solutions for obtaining higher resolution and larger collecting area than the Space Telescope are currently under consideration by NASA: The Very Large Space Telescope, The Coherent Optical Telescope Array (see paper by Gursky and Traub in this conference), and the 100 Meter Thinned Aperture Telescope (see paper by Korsch and Warner this conference).

Telescope sensitivity increases as the square of the aperture, therefore, for a tenfold increase in sensitivity over the Space Telescope (ST), an aperture of at least 7 meters is required. A segmented mirror seems inevitable since the technology for fabricating lightweight single piece mirrors above 3 to 4 meters does not exist today and is not foreseen. For this and other reasons, it is unlikely that the same nearness to diffraction limited performance as the ST can be achieved. Therefore, an aperture larger than 7 meters is desirable.

Apertures from 6 to 10 meters have been investigated (Figure 1) assuming that no single piece mirror segments larger than 3 to 4 meters will be available. Of course, minimizing the number of segments will minimize the complexity of figure control, alignment, packaging, launch, and on-orbit assembly. Studies of large aperture mirrors in space have led to deployable concepts and on-orbit assembly concepts utilizing the Space Shuttle Cargo Bay (useable diameter approximately 4 m). These concepts are very attractive at microwave or infrared wavelengths. For diffraction limited visible imaging telescopes these concepts seem less appropriate.

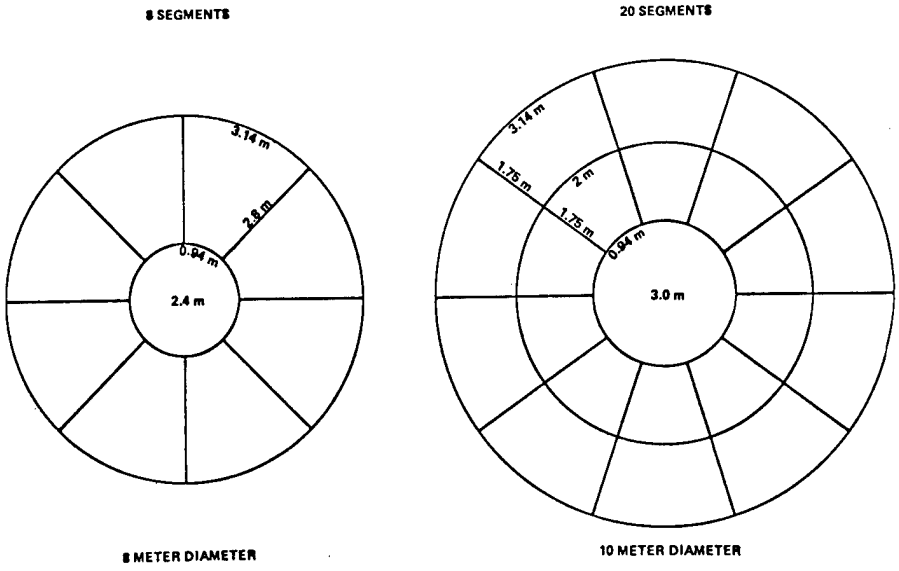


Fig. 1. Segmented primary mirror concepts

Typical scientific instrument complement

The VLST will simultaneously accommodate a variety of instruments designed in a modular fashion, in either a radial or axial configuration similar to the Space Telescope approach. The interfaces will be such that all instruments are completely interchangeable and replaceable on-orbit. It is likely that a wide field imaging camera will occupy the center field area where off-axis astigmatism is not significant.

Launch concept

Launch vehicles of the Saturn V class or the contemplated Heavy Lift Launch Vehicle (HLLV) would be required to transport a preassembled mirror larger than 4 m to orbit. However, a novel launch concept, which allows launch of an 8 m diameter preassembled mirror into orbit without requiring a new launch vehicle, is proposed in this discussion (Figure 2).

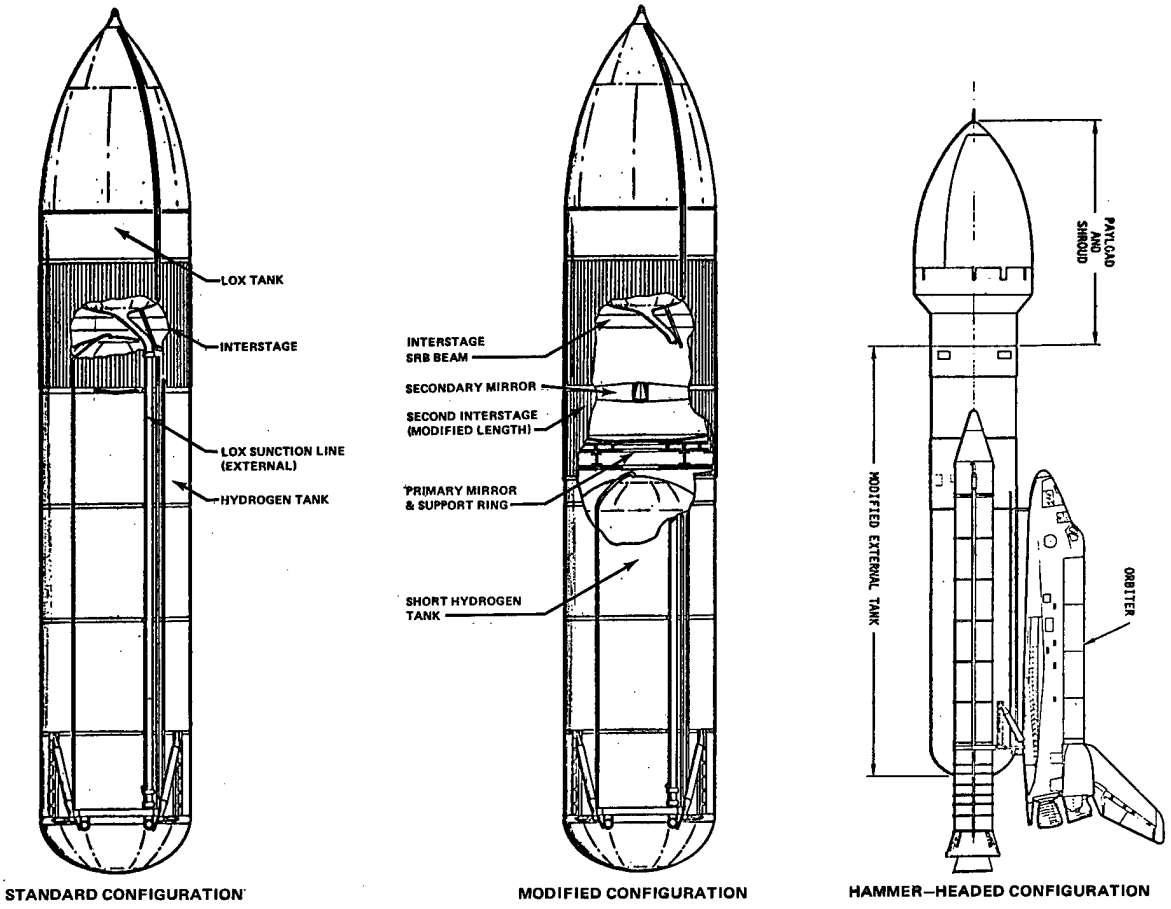


Fig. 2. VLST launch concepts in the Space Shuttle External Tank

The launch concept is to substitute a shortened hydrogen tank in the Shuttle External Tank (ET) envelope and use the leftover volume to carry the secondary mirror, support structure, and 8 m preassembled primary mirror to orbit. Orbiter total payload capability under these conditions is approximately 50% above the telescope launch weight of 25,000 lbs. Another possibility which has been studied is to hammerhead the external tank and thereby provide a ten meter diameter cargo volume. Should later analysis invalidate the use of the ET as a cargo carrier, the primary mirror will have to be transported as a folded or disassembled structure inside the orbiter bay. This would require complex on-orbit assembly and alignment which is considered a major technology problem.

Launch and on-orbit assembly phase

The preferred option for orbital construction of the 8 m aplanatic Cassegrain telescope, envisions launch of the assembled and coarse aligned primary mirror and the secondary mirror inside a modified interstage of the Shuttle ET. Figure 3 depicts the launch and assembly phases. Upon reaching orbital altitude of about 425 km the ET is vented of all re

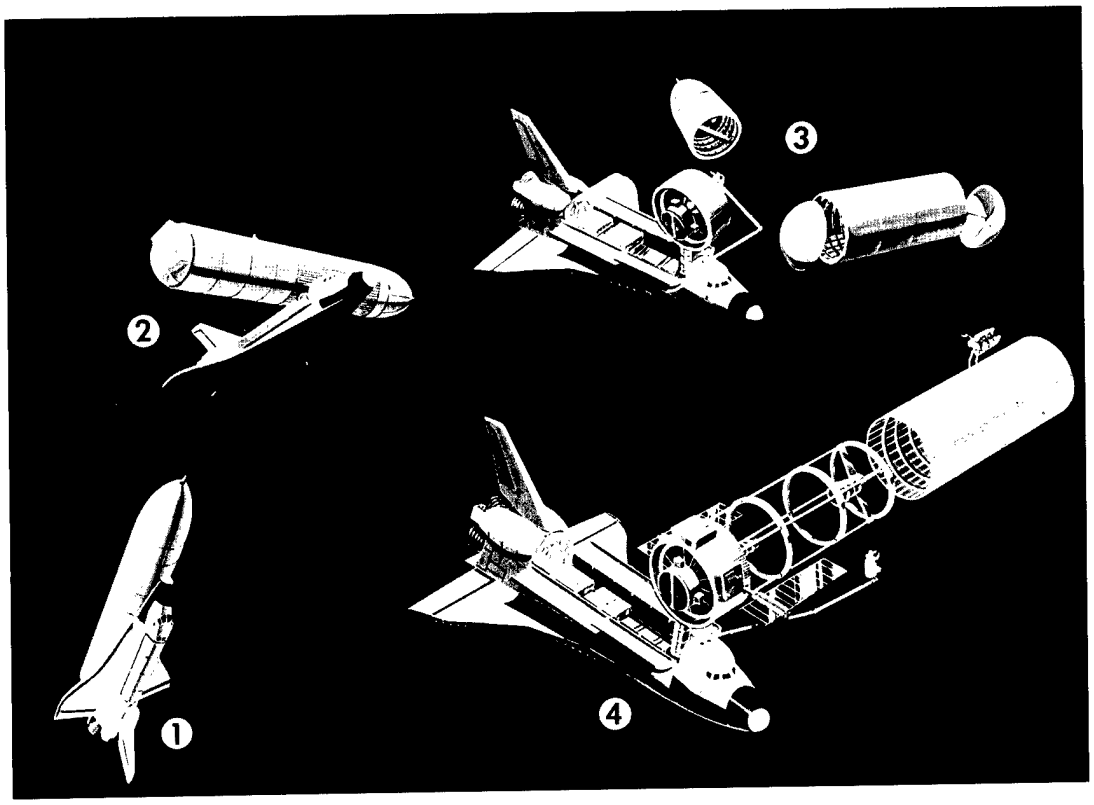


Fig. 3. Orbital assembly phases

sidual propellants. The Shuttle then separates from the ET and docks to a berthing port on the ET interstage. Astronauts using the remote manipulator system (RMS) with a work platform will then disconnect the LOX tank main interstage and forward shroud and the hydrogen tank from the modified interstage section which contains the primary mirror, mirror support structure, secondary mirror and instrument module support structure. Once the interstage, containing the telescope mirrors, has been disassembled from the tanks it becomes the telescope spacecraft structure. Power modules and stabilization modules are extracted from their launch location on Shuttle pallets via the RMS and are mounted externally to the interstage spacecraft structure. Metering rods, also carried in the Shuttle bay, are installed to mating rod sections mounted to the primary mirror support structure. The secondary mirror spider frame with its mirror is extended out from the interstage. After securing and reinforcing the secondary mirror spider and metering structure and activating the power modules and stabilization module, the as yet incomplete telescope may be "parked" in orbit and the Shuttle may return from orbit. During the next visit the Space Shuttle will carry another ET into orbit, the hydrogen tank of which has been modified in such a manner that it can be removed from the ET structure, its forward and aft bulkheads can be cut and the entire cylindrical section can then be installed to the telescope spacecraft structure, as a light shield. Further construction and assembly is then resumed through installation of the conical light baffles at the primary and secondary mirrors, installation of "checkout" instrument modules and laser interferometers for telescope alignment, leading up to preparing the telescope for operation.

Alternatively the above complex assembly procedure could be minimized through use of the compact intermediate incidence Cassegrain design (as described below) which can be launched in a more fully assembled configuration. Assembly would be limited to addition of peripheral components and final alignment.

Telescope optical design

One optical design tentatively selected is an aplanatic Cassegrain commonly referred to as a Ritchey-Chretien (Figure 4a, b). The central obscuration is estimated at 30% based on stray light baffling concerns. The basic assumptions about the design were:

(1) 8 m diameter primary
(2) F/2.2 primary
(3) Approximately 2.5 m back-focal distance

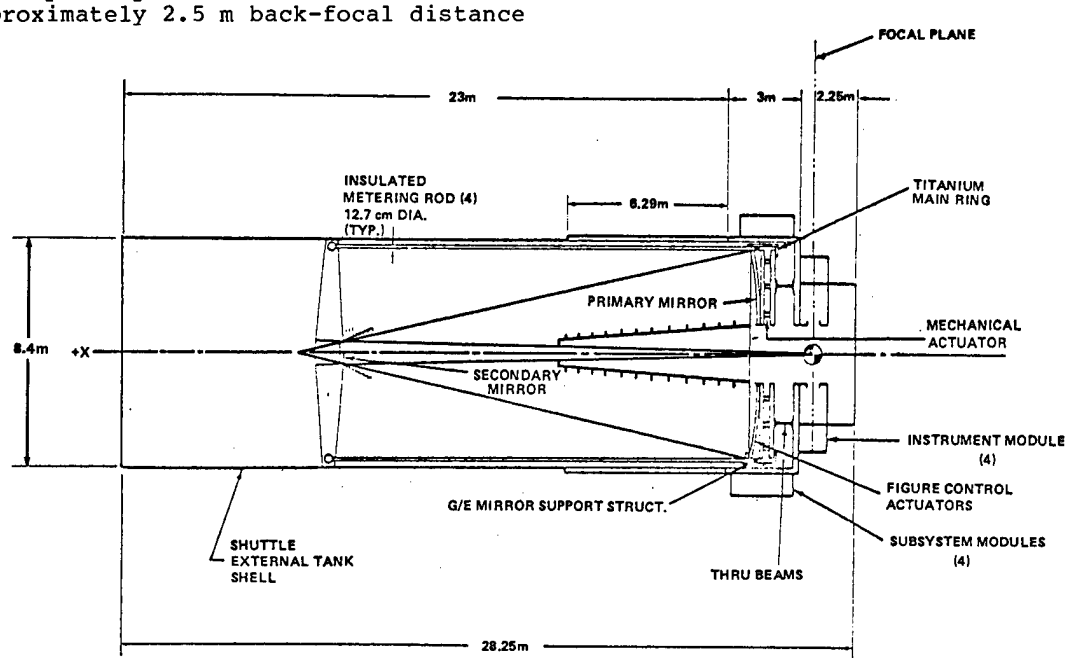


Fig. 4a. VLST Ritchey-Chretien aplanat
8 m Primary f/2.2, f/24 system

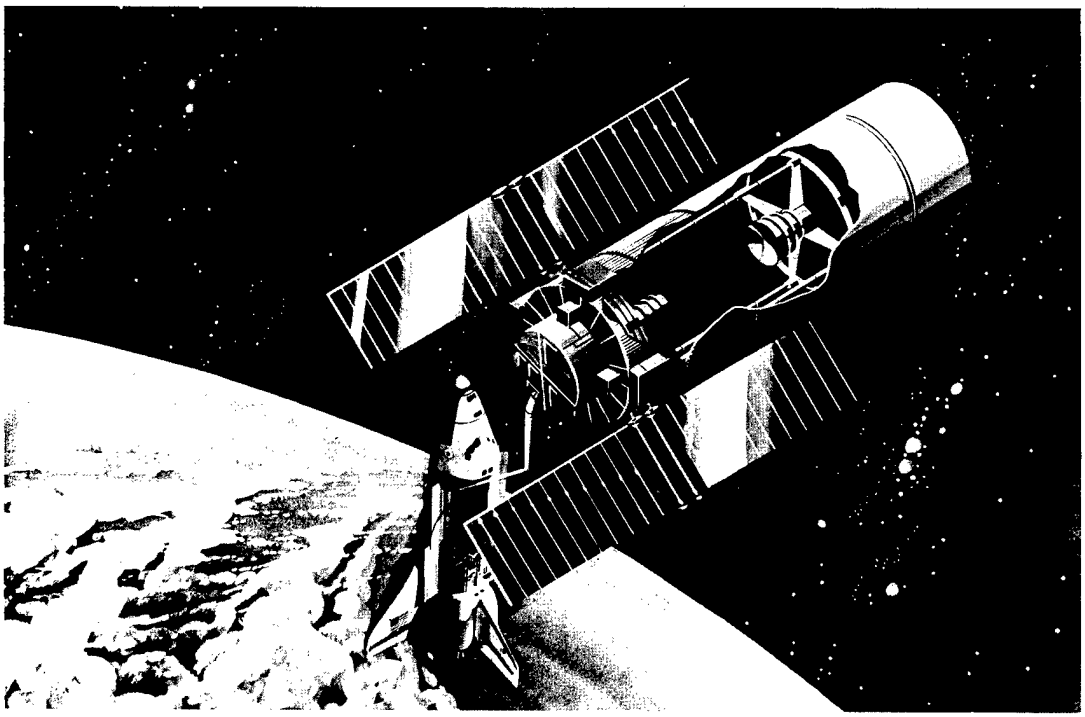


Fig. 4b. VLST during Shuttle servicing

An alternative optical design for an aplanatic Cassegrain has recently been proposed.* This new class of designs offers intermediate solutions between the near-normal incidence Ritchey-Chretien aplanats and the near-glancing incidence Wolter aplanats. Designs investigated thus far around 45 degrees (Figure 5) look particularly attractive because of their compactness. These compact designs not only reduce the vehicle moment of inertias in two axes and alleviate the pointing control system difficulties but introduce the prospect of launching a fully assembled telescope inside the ET interstage or an extension of the forward section of the ET. Another fortuitous feature of the intermediate incidence aplanat

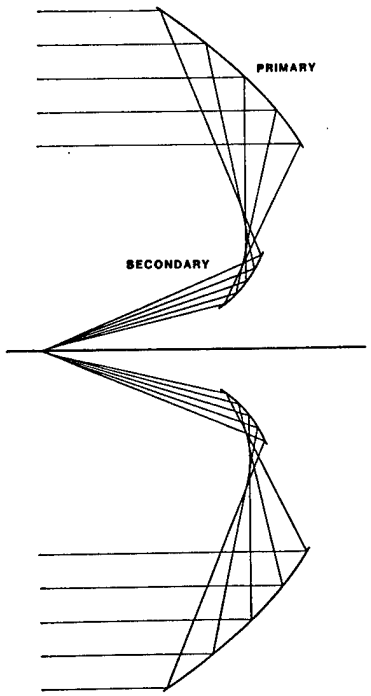


Fig. 5. VLST intermediate incidence aplanat concept

design is that the neutral point of the the secondary mirror can be adjusted to be coincident with the center-of-mass of the mirror itself. This means the secondary mirror could be rotated about its center-of-mass with no first order influence on image quality. For either image motion compensation or background chopping, moving the secondary mirror about its center-of-mass has definite advantages.

Mirror structural concepts

Although traditional lightweight mirror fabrication can achieve weight reductions of 80%, it is unlikely that it will meet the payload constraints of this VLST concept.

Newer ultra lightweight mirror structure concepts using active control show more promise for this application. These newer approaches retain a thin glass faceplate for the polished surface but employ lighter weight materials such as composites for the mirror substrate structure. Graphite-epoxy is most often used for this application although other composites are being considered (Figure 6). Associated with the composite structures are higher variability of coefficients of thermal expansion (as compared to ULE, ZERODUR or CERVIT) which make figure control a necessity.

*Korsch: Applied Optics, 19, 499 (1980).

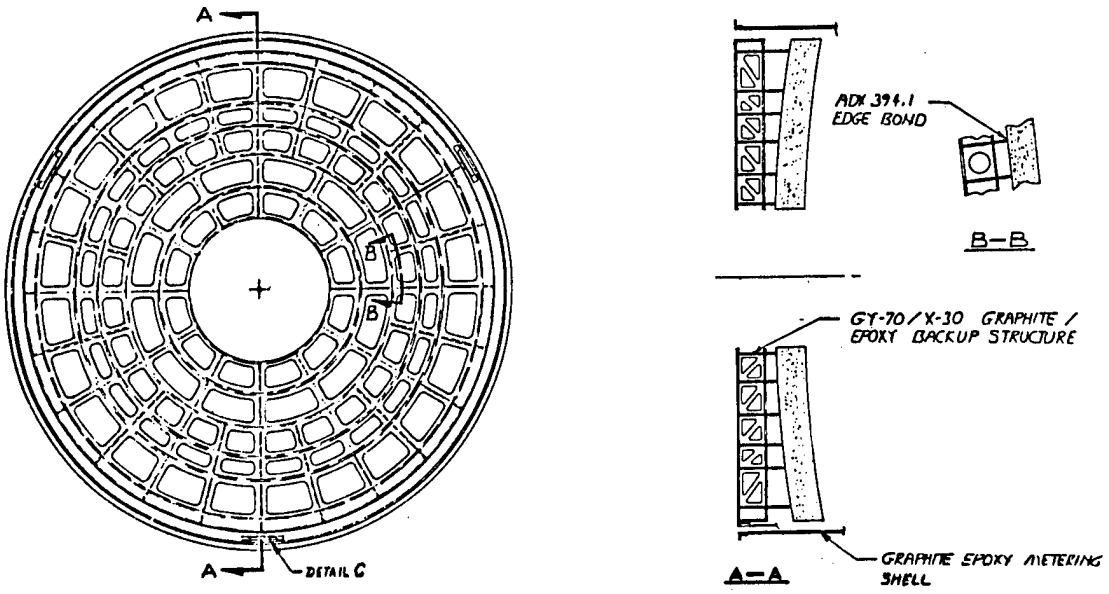


Fig. 6. Graphite epoxy structure concept for VLST primary mirror

Performance control sensors

The VLST will employ a performance control system for monitoring the image quality and tuning the telescope to peak performance. The sensors for this system will consist of white light stellar interferometers at several field points in the focal plane. By measuring the wavefront error in the exit pupil from several field points (at least two are required), the source of the error can be traced to either the figures of the individual mirrors or misalignments between the mirrors.

Figure control of flexible and/or segmented mirrors is a well established technology dating back more than 10 years. Early experimental work dealt only with static figure corrections for applications such as the Space Telescope where the mirror dynamic modes could be ignored. More recent work has been directed to high speed wavefront corrections.

Technology requirements

In order to develop a Very Large Space Telescope for the 1990 time frame, certain technology must be available in the mid-1980's to provide the basis for eventual design, development and on-orbit assembly of such a system. A first order assessment of the complex requirements for the VLST in optical technology and orbital construction are identified vis-a-vis the current state-of-the-art in Table 1.

The driving technologies include lightweight mirrors, polishing off-axis paraboloids and optical diagnostic sensing and controls. Lightweight mirror technology ≤ 25 kg/m^2 is now available, but only for small mirror sizes. Mirror segments up to four meters must be capable of being produced in periods of time of a year or less. Some recent success has occurred with automated polishing of large mirrors leading to the hope that off-axis paraboloids of substantial size may be treated in the near future. Our current state-of-the-art in wavefront sensing and control is embodied in the static system of Space Telescope. This must be evolved into one of continuous realignment and control.

Table 1. Prime technology requirements

	CURRENT STATE-OF-THE-ART	VLST REQUIREMENTS
MIRRORS	ST – 76% LIGHTWEIGHTED ULE APPROXIMATELY 166 kg/m^2	25 – 30 kg/m^2
COATINGS	2.4 m APERTURE UV – COATING ON GROUND	3 m APERTURE UV CLEANING & RECOATING ON ORBIT
POLISHING	2.4 m DIAMETER ST PRIMARY NEAR DIFFRACTION LIMIT λ/65 RMS	3 m PIE – SHAPED SEGMENTS OFF-AXIS FIGURE λ/65 RMS
DETECTOR	25 mm SQUARE CCD – 25 μ/PIXEL FILM RESOLUTION APPROXIMATELY 1 μ	>50 mm SQUARE CCD – 10 μ/PIXEL INCREASED FILM RESOLUTION – <1 μ ON – ORBIT STORAGE AND HANDLING
ORBITAL ASSEMBLY, FABRICATION	CUTTING, JOINING OF STRUCTURAL COMPONENTS, DEPLOYMENT OF SOLAR SHIELD (SKYLAB), BEAM MANUFACTURING (SPACELAB)	ALIGNMENT CONTROL, CUTTING JOINING PRECISION POSITIONING OF COMPONENTS COATING & PAINTING IN SPACE
OPTICAL DIAGNOSTIC SENSORS & CONTROLS	STATIC, INTERMITTENT RE-ALIGNMENT AND FIGURE CONTROL	CONTINUAL REALIGNMENT AND FIGURE CONTROL WAVE FRONT ERRORS <0.025 @ 633 nm
POINTING STABILITY	0.007 ARC SECOND RMS	0.002 ARC SECOND RMS

One of the options for increased sensitivity and angular resolution of the Very Large Space Telescope is probably the most straightforward: extension of the Space Telescope design, with much of the VLST technology directly dependent on that of ST. This would provide the basis for a cost effective extension of the current study efforts for subsequent application to a larger, more complex format. On the other hand, the increases in resolution and collecting area, although providing a substantial capability, are generally small compared with those proposed for the Coherent Optical Telescope Array or the 100 Meter Thinned Aperture Telescope.

Acknowledgement

We gratefully acknowledge the advice and support provided by C. R. O'Dell in deriving the VLST concept.

Bibilography

1. Berggren, R. R. and Lenertz, G. E., Feasibility of a 30-Meter Space Based Laser Transmitter, ITEK Report, NASA Contract NAS3-19400, October 1975.

2. Golden, L. J., Shack, R. V., Slater, P. N., Study of an Instrument for Sensing Errors in a Telescope Wavefront, Final Report NASA Contract NAS8-27863, Optical Sciences Center, University of Arizona, May 1974.

3. Facey, T. A., Focal Plane Wavefront Sensor Definition Study for Space Telescope, Final Report NASA Contract NAS8-30351, Perkin-Elmer Corporation, September 1976.

4. Technology Study for a Large Orbiting Telescope, Final Report NASA Contract NASw-1925, ITEK Corporation, May 1970.

5. Study of Telescope Maintenance and Updating in Orbit, ITEK Report 68-8599-1, NASA Grant NGR-31-001-004, ITEK Corporation, May 1968.

6. Space Telescope Phase B Definition Study, Final Report NASA Contract NAS8-29948, Perkin-Elmer Corporation, March 1976.

7. Advanced Shuttle Projects, Rockwell International ASP-78-087, March 1978.

8. Korsch, D., Aplanatic Two Mirror Telescope from Near Normal to Grazing Incidence, Appl. Opt., 19, 499 (1980).

Coherent optical arrays for space astronomy

W. A. Traub, H. Gursky
Harvard/Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, Massachusetts 02138

Abstract

The conceptual basis of an orbiting observatory designed for high angular resolution optical astronomy is presented. Operating above the atmosphere, an array with a 15 meter baseline could achieve a one-dimensional angular resolution of 0.005 arc sec in the visible. Broad-band detection from the UV to the near infrared should be possible. To obtain two-dimensional images at high resolution we expect to employ tomographic reconstruction techniques, followed by sidelobe removal if necessary. Active optical elements will be needed in order to maintain coherence of beams in the image plane. The instrument is an optical analog to the Very Large Array in radio astronomy and a space analog to the Multiple Mirror Telescope. Prototype designs consisting of four to six mirrors are discussed for the case of a single shuttle payload.

Introduction

As Space Telescope comes closer to realization, it is becoming increasingly clear that further advances in optical astronomy will critically depend upon our ability to transcend the present barriers of angular resolution and collecting area that are imposed upon us by single-element primary mirrors. The diffraction-limited beam size of a telescope not only prevents one from observing phenomena on a finer angular scale, but also ultimately dictates the limiting magnitude that can be achieved, solely because of the background sky brightness. In this article we will discuss the technical aspects (but not the scientific rationale) regarding very high angular resolution telescopes, in part because the first steps in this direction have yet to be accomplished, namely the leap from ground-based imaging ($\sim$ 1 arc sec), to ST imaging ($\sim$ 0.1 arc sec). We propose[1,2] to extend this angular resolution capability by more than one more order of magnitude, down to the $\sim$ 0.005 arc sec level, while still maintaining a large collecting area, and with options to push both parameters still further.

Optical concepts

The concept of an array of coherent optical telescopes can be considered as the optical analog to the instruments that radio astronomers have been building and using for many years now, namely very long baseline interferometers (VLBI), and the Very Large Array (VLA). The main difference occurs in the photon detection process, since in the radio region separate coherent detectors can be assigned to individual antennas, whereas in the optical region we generally combine the beams from each collecting mirror in a coherent fashion before the light can be detected. Bearing this in mind, we refer to Fig. 1 which shows an optical schematic design for six telescopes arranged in a linear array, with a combined focal plane at the center of the array. Here each telescope has a primary mirror diameter of 2.4 m, the same as for ST. The packing and beam-folding are chosen so as to fit the ensemble into the Shuttle bay (18.3 m length, 4.6 m diameter). The requirement of coherent beam recombination dictates that the fold mirrors form equal path optical delay lines for each beam.

The beam pattern of such an array is of course substantially narrower along the array axis than it is in the perpendicular direction, as we show in Fig. 2. Here we see that the ideal diffraction pattern has a cross-section of about 0.005 arc sec by 0.05 arc sec. Point sources would thus appear in the focal plane as narrow ovals with aspect ratio of about 10 to 1.

If we consider the images formed by this array in one orientation compared to those formed when the array is rotated up to one-half turn about its optical axis, it then becomes clear that we now have available all of the information needed to construct a fully two-dimensional image at the higher resolution of 0.005 arc sec. The mathematical technique is closely related to the tomographic methods currently being employed routinely in medical CAT scanners.

There are many possible optical configurations. An alternative design is shown in Fig. 3 where we have used four 1 m diameter mirrors spaced on a 10 m baseline. This is the prototype of a "minimum redundancy array", so-called in the radio astronomy literature because of its property of yielding six independent, unique mirror spacings, being respectively 1,

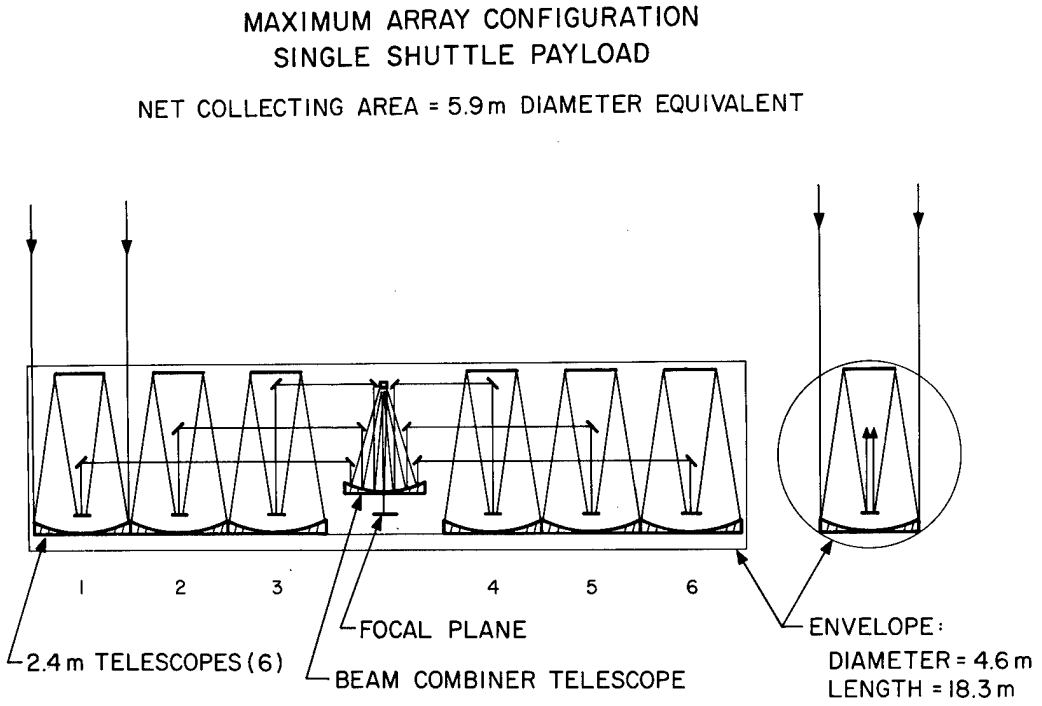


Fig. 1. Optical schematic for a 6 mirror array, using individual primaries which are each 2.4 m in diameter (i.e., the same as ST). This is about the maximum amount of collecting area that can be accommodated in a single shuttle payload.

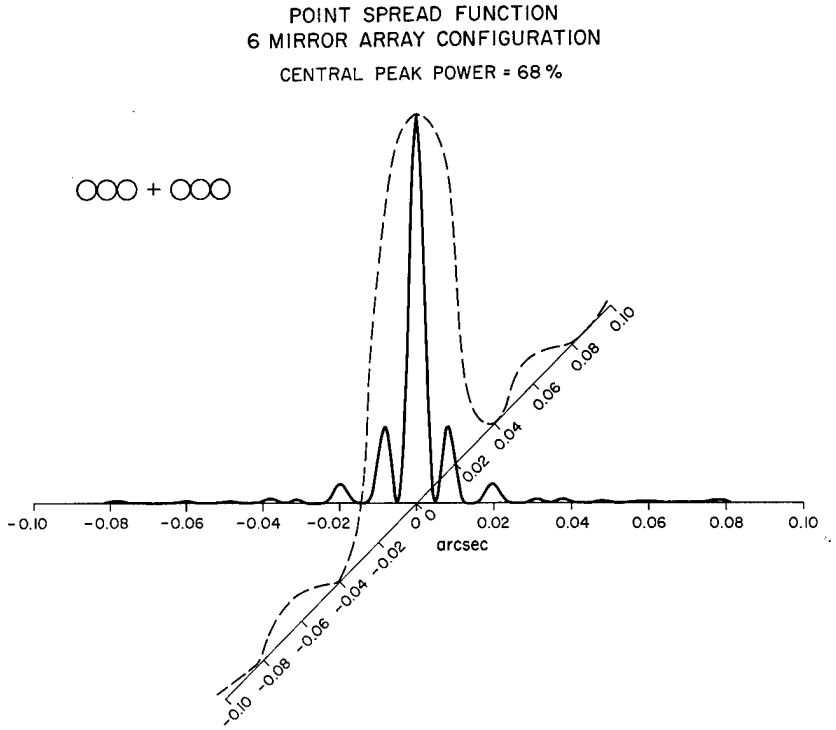


Fig. 2. Diffraction pattern for the array shown in Fig. 1. The central lobe has a width of about 0.005 arc sec in the narrow direction, and about 0.05 arc sec in the orthogonal.

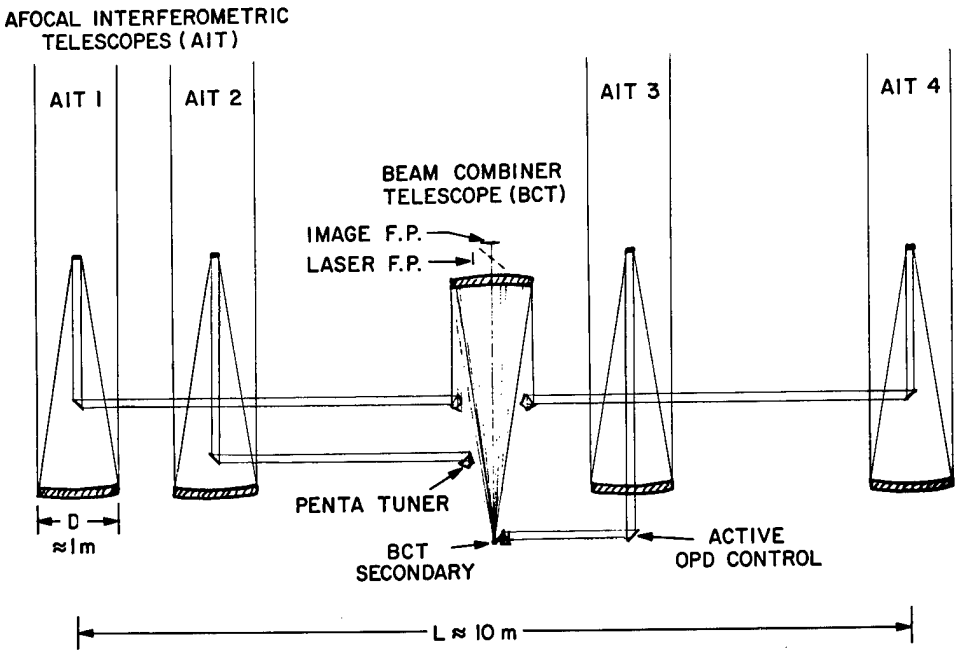


Fig. 3. Optical schematic for a 4 mirror system, arranged as a minimum redundancy array. Chosen for illustration purposes, these mirrors are 1 m diameter on a 10 m baseline.

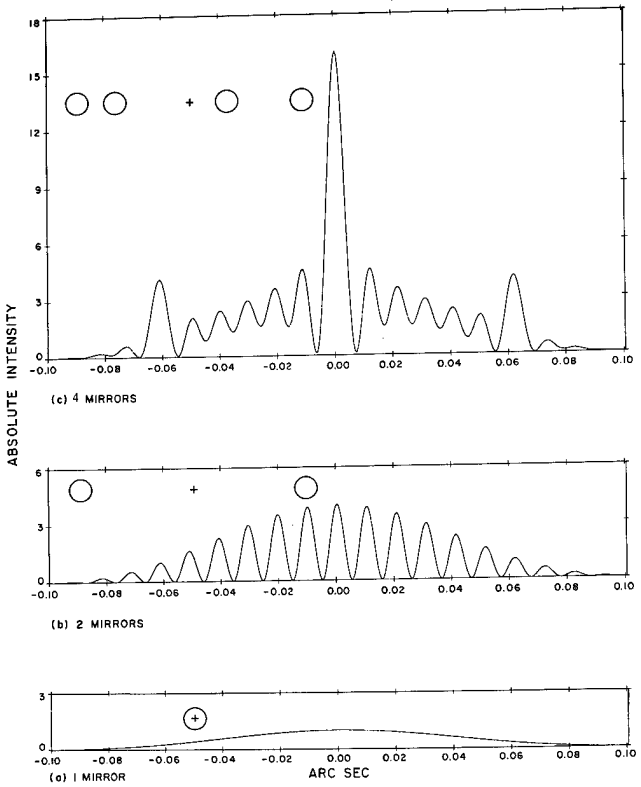


Fig. 4. Diffraction pattern in one dimension for the following cases: (a) a single 1 m diameter mirror; (b) two 1 m mirrors on a 10 m baseline; and (c) four 1 m mirrors, as in Fig. 3.

2, 3, 4, 5, and 6 times the distance between the nearest spaced pair. This property also yields equally spaced sampling points in the angular transform domain. The one-dimensional diffraction pattern of this array is shown in Fig. 4, where it is also compared to the patterns which arise from a pair of mirrors and a single mirror, respectively, plotted to the same scale.

Arrays such as these can in principle reach very faint magnitude limits in reasonable integration times: assuming 5 percent overall efficiency, about 180 events/hour in the central fringe, and a 1000 Å bandwidth, the array in Fig. 1 will reach V=28, and the array in Fig. 3 will go to about 25th visual magnitude.

It is also possible to design these arrays so that the focal plane appears at one end of the chain, rather than in the middle; in this case one can join two or more such arrays to a single focal plane, thus allowing the dual options of increasing the angular resolution by a factor of two or more in one dimension, and/or building in the orthogonal direction to achieve better two-dimensional imaging. In either case, the in-orbit assembly requirements would be minimal.

Advanced technology requirements

Coherent telescope arrays of the type being proposed here clearly would benefit from the utilization of a number of advanced technologies; the arrays may in fact totally depend upon such technologies. We discuss here a few salient areas in which we believe some degrees of development will be required.

Active optics

There are typically about 7 reflections per beam in the arrays under discussion, so the wavefront distortion at each reflection must be held to a minimum in order to ensure that the final image quality is not substantially degraded. Certainly the nature and scale size of polishing irregularities will eventually be a consideration, but if for the moment we ask that the peak-to-peak final wavefront distortion be less than $\lambda/4$, it then follows that most of the optical elements will have to have surfaces good to approximately $\lambda/100$ (rms), depending on the additive nature of the errors. Clearly there is a potential role here for some sort of wavefront modification device in each of the telescope beams. The active component need not have a fast response, as the errors are likely to be changing slowly, on a scale of minutes. Possibly one could employ a technique similar to that used on the ST, where figure correction of the primary mirror is effected by 24 pistons on its back plane. Sensing of the wavefront can be done by examination of a bright star in the focal plane, either iteratively with all mirrors or pairwise, with the other beams blocked.

A second requirement in the active optics system is focus, beam steering, and path length control. The question of how to sense these quantities is still open, although the actuating can clearly be accomplished to a small fraction of a wavelength if one is allowed a sufficient number of laser beam probes. Here again it is conceivable that all these functions could be sensed by simply analyzing the focal plane image quality, and making either continuous or occasional corrections. An alternative is to employ a small number of carefully chosen paths for a laser beam sensing system, and to turn on the sense/adjust system at appropriate intervals.

Structure

A passive, stable structure will greatly reduce the dynamic range requirements on the active optics system. Assuming the availability of a high quality composite material such as graphite epoxy (or graphite magnesium), we will have a thermal coefficient of 10^{-7}/°C, giving a length change of 1 micron/°C over a 10 m baseline. With thermal controls to keep the temperature constant to about 1°C, the active system then has a comfortably small range over which it must work.

Aspect

Sensing and control of the optical axis with respect to inertial space is a critical factor: with a field of view of roughly 3 arc sec, and a pixel size of about 0.003 arc sec, we not only need to acquire an object and maintain it in the approximate center of the field, but we must also ensure that the pointing vector variation stays within a pixel between successive focal plane read-out cycles. It is not clear that gyros alone will be able to provide the required accuracy. It may be that aspect drift sensing can be accomplished by examination of the focal plane light distribution, or by auxiliary detectors located over a larger field adjacent to the central axis. For a linear array the instrument must be allowed to take on all possible azimuth positions in order to provide the information needed to reconstruct the image in two dimensions; spin rates on the order of one revolution per orbit may be adequate.

Detectors

Photon counting imaging detectors are rapidly becoming state-of-the-art. For typical focal lengths, pixel sizes on the order of 10 to 30 microns are adequate, but array sizes of at least 1000 by 1000 elements will be needed for most astronomical observations. These detectors are in general ideally suited to low-light-level applications, with dark-count rates of about 1 count/pixel/hour for currently available systems. Output signals in the form of photon arrival coordinates, plus time, are ideally suited to computer processing of data. We also note that a slit spectrograph would be well-matched to the narrow image produced by a linear array.

Conclusion

We have described in this paper the conceptual basis of an optically coherent telescope array in space, and have begun to identify those advanced technologies which will be required. The scientific advantage of a telescope array with respect to Space Telescope is based on its potential for achieving large increases in both angular resolution and collecting area. Cost control can be realized through the use of identical, modest-diameter optical elements arranged in a modular fashion. We are now in the process of beginning to study in greater detail these and other aspects of the array concept.

Acknowledgements

We have drawn on resources both at the Itek Corporation and Marshall Space Flight Center in the preparation of this paper. We also gratefully acknowledge discussions with G. Colombo and N. P. Carleton, both at the Harvard/Smithsonian Center for Astrophysics.

References

1. Gursky, H. and Traub, W. A., "Use of Coherent Arrays for Optical Astronomy in Space," in Space Optics, SPIE Vol. 183, pp. 188-197, 1979.

2. Traub, W. A. and Gursky, H., "Coherent Arrays for Optical Astronomy in Space," in Telescopes for the 1990's, to be published, 1980.

Super-size space telescope

Dietrich Korsch
TAI Corporation, Huntsville, Alabama 35803
John W. Warner
Marshall Space Flight Center, Alabama 35812

Abstract

A class of space telescopes for astronomical observations with a resolution and collecting capability more than one order of magnitude better than what is expected from the 2.4 m Space Telescope is discussed. To this purpose aplanatic two-mirror systems of coplanar primary/secondary mirror arrangements with approximately 45° angles of incidence and an overall diameter of about 100 m have been designed and analyzed. The main advantages of these systems are their compactness and the associated minimization of the moment of inertia in two axes. Two opposing secondary arrangements, one forward-reflecting and the other backward-reflecting are analyzed and compared.

Introduction

What astronomy will be in the year 2000, and what the significant areas of research will be, is difficult to predict in any detail since the doubling time for information is about 10 years. However, as in the past, most major advances in the field have depended on increases in aperture size and detector sensitivity. Notable examples include the transition from the use of the human eye to film and thence to photoelectric devices as well as the building of the Hale 200 inch telescope. It is expected that future astronomers will regard the era of the Space Telescope as a major milestone in terms of increased resolution and sensitivity. Its 2.4-m mirror and complement of five scientific instruments will be able to resolve sources 10 times smaller and detect stellar sources 50 times fainter than is possible with groundbased telescopes.

Yet the era of Space Telescope will itself reveal questions and pose unsolved problems which will require new capabilities for astronomical observation. Even though these questions are as yet unknown, new concepts for increasing aperture and resolution can be studied which will address major classes of problems involving the study of the planets, stars, galaxies and the universe as a whole. One of these concepts being studied by Marshall Space Flight Center is the 100-m Thinned Aperture Telescope (TAT).[1, 2]

Design considerations

The larger the space telescope, the more decisive become mechanical and structural considerations. To make the maintenance of alignment manageable, and to alleviate slue and pointing maneuvers, it is imperative to compress the system as much as possible, thereby minimizing the moments of inertia. It is obvious that if the collecting area is to be maximized, structural compression can only be achieved along the dimension orthogonal to the plane of the entrance aperture.

The simplest optical system would be a single paraboloid. However, even at a focal ratio of one the focal distance for a 100 m aperture would also be 100 m. We shall, therefore, concentrate our study on two-mirror configurations, folded for maximum compactness. An artist's concept of one such configuration is shown in Figure 1. The drawing also indicates that the aperture is only partially filled using segmented elements of mirrors. This has the advantage that initially the aperture may be equipped with relatively few mirror elements, the number and arrangement of which would have to be determined to yield an adequate diffraction pattern. To maintain alignment and the surface figure, each element must be actively controlled.

Aplanatization

Since it is desirable for the telescope to provide a diffraction limited image, even though only over a relatively small field, it may be necessary to aplanatize the system. Surface equations for aplanatic two-mirror telescopes have been developed and published.[3] These equations, however, are very complex and the experience taught that a numerical evaluation of the differential surface equations which are of much simpler form is a more convenient procedure. We shall give here, therefore, a brief derivation of the two differential equations, using Figure 2.

The basic design parameters are:

ρ_{01} = center radius of the first element

ρ_{02} = center radius of the second element

d = center to center distance

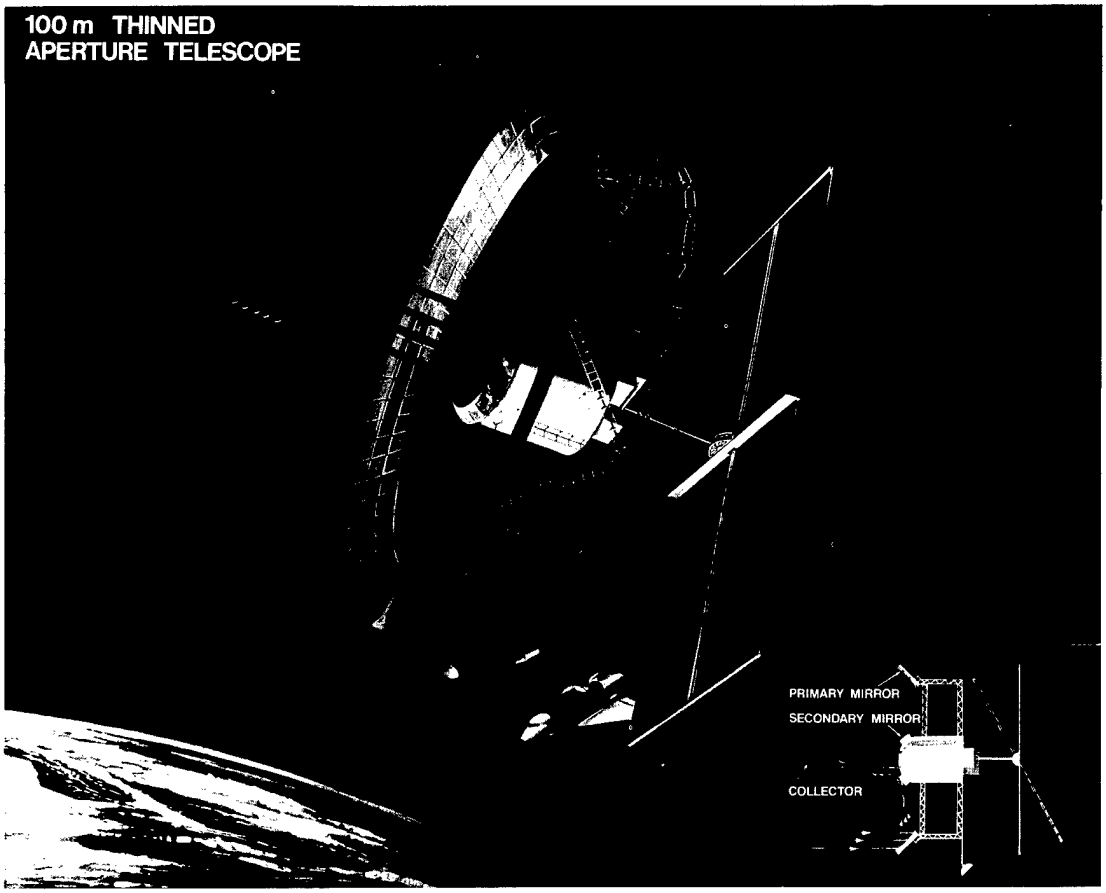


Figure 1. Artist's concept of a large aperture space telescope.

b = back focal distance, measured from the center of the second element.

The two relations defining aplanatism are:

$$s_1 + s_2 + z_1 = c \tag{1}$$

with $c = [(\rho_{01} - \rho_{02})^2 + d^2]^{\frac{1}{2}} + [\rho_{02}^2 + b^2]^{\frac{1}{2}}$ for the absence of spherical aberration, and Abbe's sine-condition for telescopes,

$$\rho_1 = f \cdot \sin\alpha \ , \tag{2}$$

where f is the system focal length.

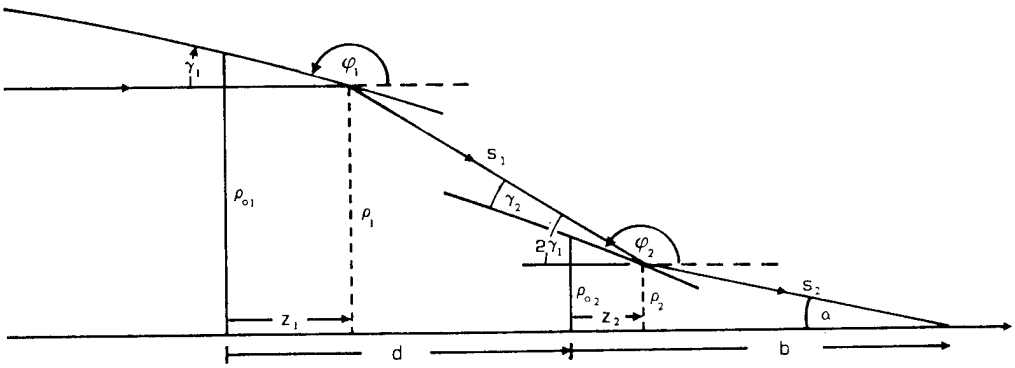


Figure 2. Parameters of the two-mirror configuration.

First surface

The differential equation for the first surface is to be of the form

$$\frac{\partial z_1}{\partial \rho_1} = \cot \varphi_1 = \cot (2\,\Pi - \gamma_1) = -1/\tan \gamma_1 \quad . \tag{3}$$

To express tan γ_1 as a function of ρ_1 and z_1, we start with the following two geometric relations derived by aid of Figure 2:

$$s_1 \cos 2\,\gamma_1 + s_2 \cos \alpha = d + b - z_1 \quad , \tag{4}$$

and

$$s_1 \sin 2\,\gamma_1 + s_2 \sin \alpha = \rho_1 \quad . \tag{5}$$

After substituting s_2 in equations (4) and (5) using equation (1), one obtains

$$s_1 \cos 2\,\gamma_1 + (c - z_1 - s_1) \cos \alpha = d + b - z_1 \tag{6}$$

and

$$s_1 \sin 2\,\gamma_1 + (c - z_1 - s_1) \sin \alpha = \rho_1 \quad . \tag{7}$$

Setting $(c - z_1) \cos \alpha - d - b + z_1 = V$,

and

$$(c - z_1) \sin \alpha - \rho_1 = W,$$

and eliminating s_1 by combining equations (6) and (7) yields

$$V (\sin 2\,\gamma_1 - \sin \alpha) - W (\cos 2\,\gamma_1 - \cos \alpha) = 0 \quad . \tag{8}$$

Using the trigonometric identities

$$\sin 2\,\gamma_1 = 2 \tan \gamma_1/(1 - \tan^2\gamma_1)$$

and

$$\cos 2\,\gamma_1 = (1 - \tan^2\gamma_1)/(1 + \tan^2\gamma_1) \quad ,$$

one obtains a quadratic equation for tan γ_1 the solution of which reads:

$$\tan \gamma_1 = -\frac{V \pm (V \cos \alpha + W \sin \alpha)}{W - (V \sin \alpha - W \cos \alpha)} \tag{9}$$

after replacing V and W again by the original variables and sin α by equation (2) we finally obtain

$$\frac{\partial z_1}{\partial \rho_1} = \rho_1 \frac{c + d + b - f\,(1 + \cos \alpha) - 2\,z_1}{(c - d - b) \cdot f \cdot (1 + \cos \alpha) - \rho_1^{\,2}} \tag{10}$$

with

$$\cos \alpha = \pm(1 - \rho_1^{\,2}/f^2)^{\frac{1}{2}} \quad . \tag{11}$$

A numerical evaluation proved the positive sign in equation (9) to be right and the sign of equation (11) follows the sign of the back focal length (positive when showing to the right).

Second surface

The differential equation describing the second surface will be of the form

$$\frac{\partial z_2}{\partial \rho_2} = \cot \varphi_2 = -\cot(\alpha/2 + \gamma_1) = \frac{\tan \alpha/2 \cdot \tan \gamma_1 - 1}{\tan \alpha/2 + \tan \gamma_1} \quad . \tag{12}$$

The task now is to write tan $\alpha/2$ and tan γ_1 as functions of ρ_2 and z_2. We start with two geometric relations which we find using Figure 2:

$$s_1 = (\rho_1 - \rho_2)/\sin 2\gamma_1 \quad , \tag{13}$$

and

$$s_2 = \rho_2/\sin \alpha \quad . \tag{14}$$

Inserting equations (13) and (14) into equation (1) gives

$$z_1 = c - (\rho_1 - \rho_2)/\sin 2\gamma_1 - \rho_2/\sin \alpha \quad . \tag{15}$$

To eliminate z_1 we use another relation that can be taken from Figure 2,

$$\tan 2\gamma_1 = (\rho_1 - \rho_2)/(d - z_1 + z_2) \quad , \tag{16}$$

and combine it with equation (15). Solved for tan γ_1 the new equation then reads

$$\tan \gamma_1 = (c - d - z_2 - \rho_2/\sin \alpha)/(f \sin \alpha - \rho_2) \quad . \tag{17}$$

Inserting equation (17) into equation (12) finally yields

$$\frac{\partial z_2}{\partial \rho_2} = \frac{\tan \alpha/2\,(c - d - z_2 - \rho_2/\sin \alpha) - (f \cdot \sin \alpha - \rho_2)}{\tan \alpha/2\,(f \cdot \sin \alpha - \rho_2) + (c - d - z_2 - \rho_2/\sin \alpha)} \tag{18}$$

with

$$\tan \alpha/2 = \rho_2/[(\rho_2^2 + (b - z_2)^2)^{\frac{1}{2}} + b - z_2] \quad ,$$

and

$$\sin \alpha = \rho_2/[\rho_2^2 + (b - z_2)^2]^{\frac{1}{2}} \quad .$$

Design examples

Three configurations have been designed and compared with each other and with an F/1 paraboloid. Each configuration was analyzed two ways, as a classical system, a combination of two cartesian surfaces which are in the case of mirrors always conic sections of revolution, and as strict aplanats.

Configuration 1:

The basic parameters are

$\rho_{01} = 40$ m

$\rho_{02} = 10$ m

$d = 0$

$b = +30$ m

$f = 126.5$ m

Entrance annulus = $(\rho_{01} \pm 10)$ m.

A semimeridional section of the aplanatic version is shown in Figure 3. A performance comparison between the classical and the aplanatic version in terms of off-axis abberations in the gaussian focal plane is shown in Figure 4.

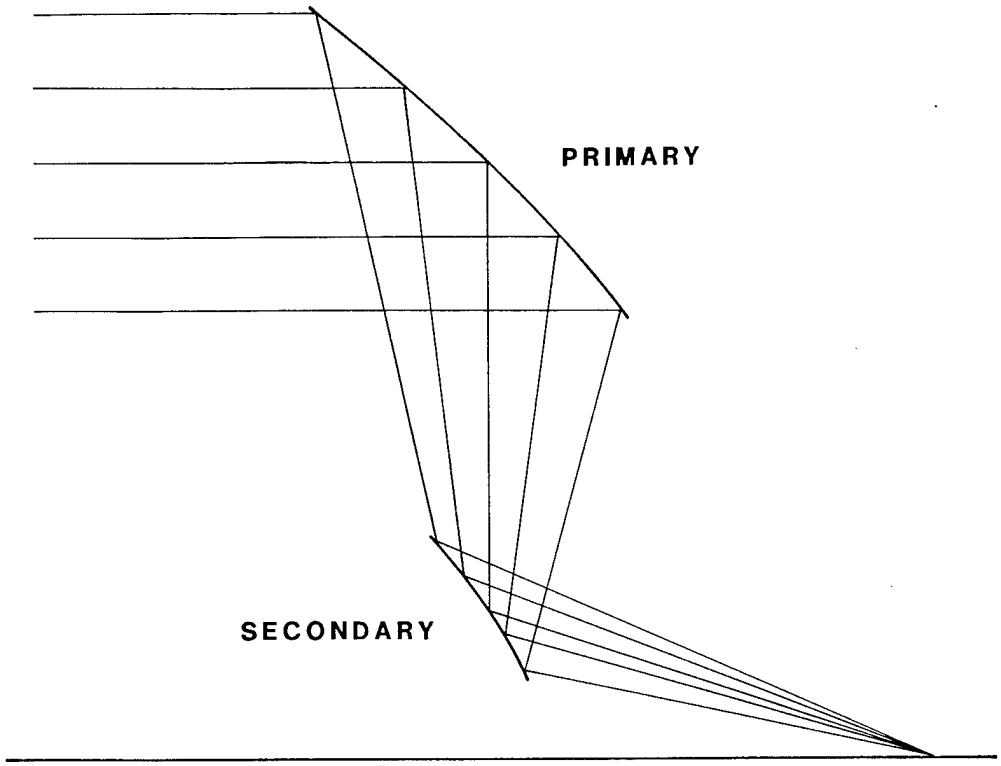


Figure 3. Aplanatic version of configuration 1.

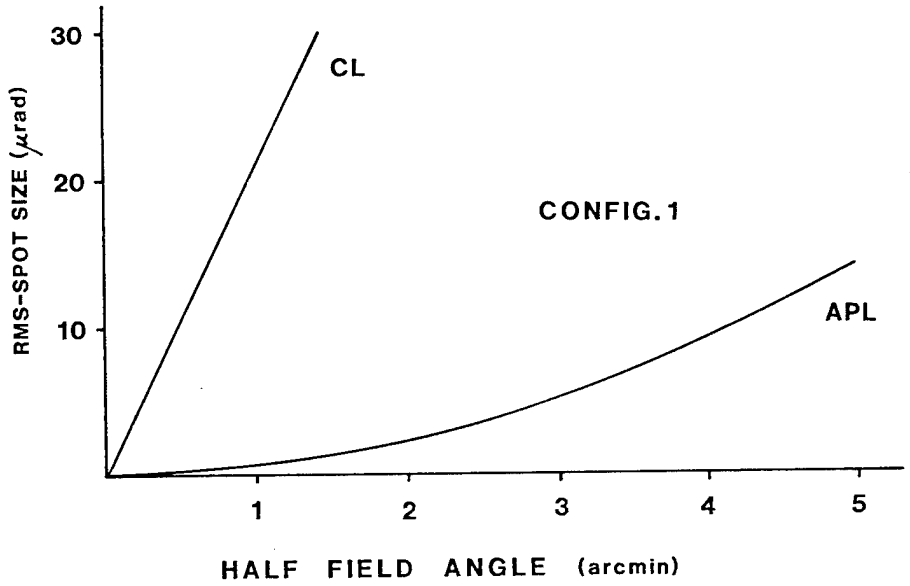


Figure 4. Performance comparison of classical (CL) and aplanatic (APL) versions of configuration 1.

Configuration 2:

The basic parameters are

ρ_{01} = 40 m

ρ_{02} = 10 m

d = -15 m

b = +30 m

f = 126.5 m

Entrance annulus = (ρ_{01} ± 10) m .

A semimeridional section of the aplanatic version is shown in Figure 5. This is the approximate geometry of a system where the common center of gravity of the secondary mirror and detector module falls into the primary mirror plane. A performance comparison between classical and aplanatic versions is shown in Figure 6.

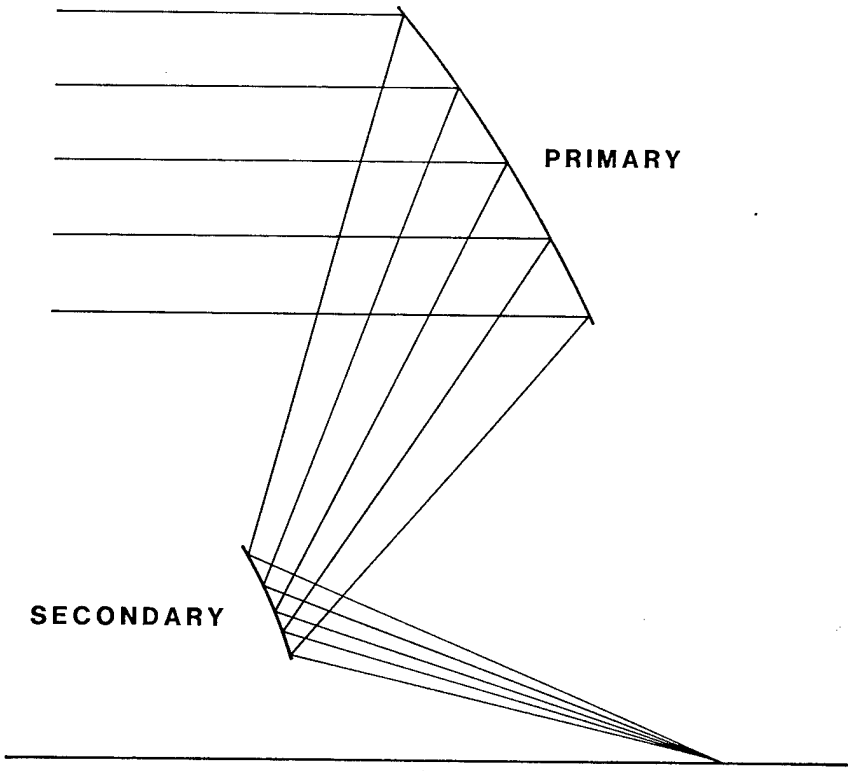


Figure 5. Aplanatic version of configuration 2.

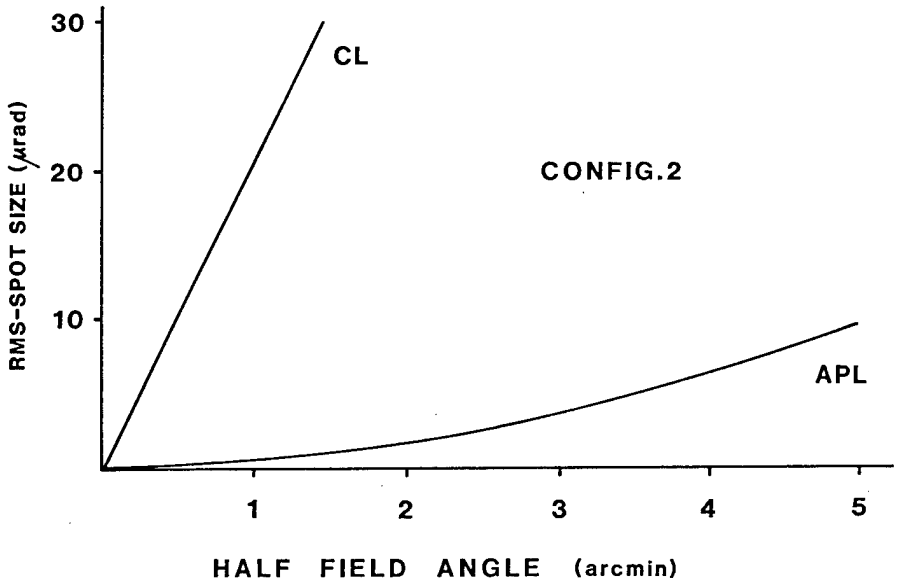


Figure 6. Performance comparison of classical (CL) and aplanatic (APL) versions of configuration 2.

Configuration 3:

The basic parameters are

ρ_{01} = 40 m

ρ_{02} = 10 m

d = 0

b = -30 m

f = 126.5 m

Entrance annulus = (ρ_{01} ± 10) m .

A semimeridional section of the aplanatic version is shown in Figure 7. An advantage of the configuration could be its potential for an effective stray light rejection. A stop could be placed at the waist of the caustic between primary and secondary which does not exist in the classical version. Also, the fact that the detector module is positioned against the incoming light might be beneficial with respect to the suppression of stray light. A comparison of the classical version which produces an excessive amount of coma, and the aplanatic version is given in Figure 8.

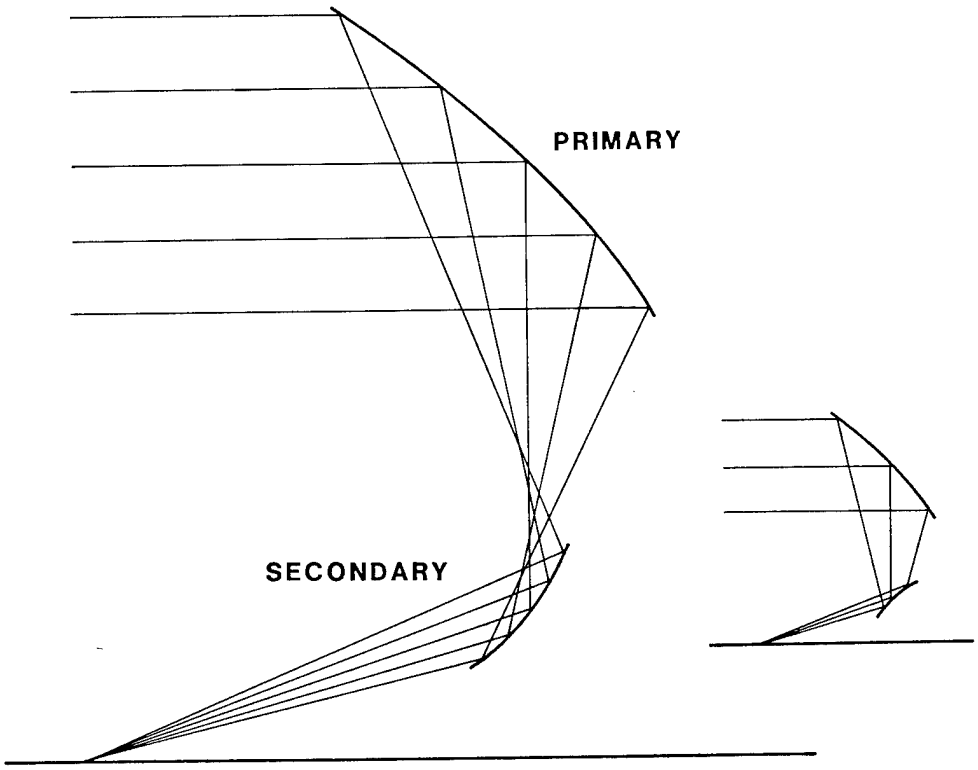


Figure 7. Aplanatic version of configuration 3. Inserted is the geometry of the classical configuration.

A comparison of the three aplanatic two-mirror systems with the field performance of a single F/1 paraboloid is shown in Figure 9.

Conclusions

In search of possible candidate designs for a very large aperture space telescope, three very compact two-mirror configurations were analyzed and compared with an F/1 single paraboloid. It seems that the off-axis performances of the paraboloid as well as of the classical versions of the two-mirror systems are rather inadequate. Excessive coma prevents a diffraction-limited performance even over a limited field. Aplanatization of the system is, therefore, desirable. The advantage of all three two-mirror configurations is their short dimension in the direction of the optical axis which minimizes the moments of inertia in two axes. Further tradeoff studies to optimize the final configuration will be necessary.

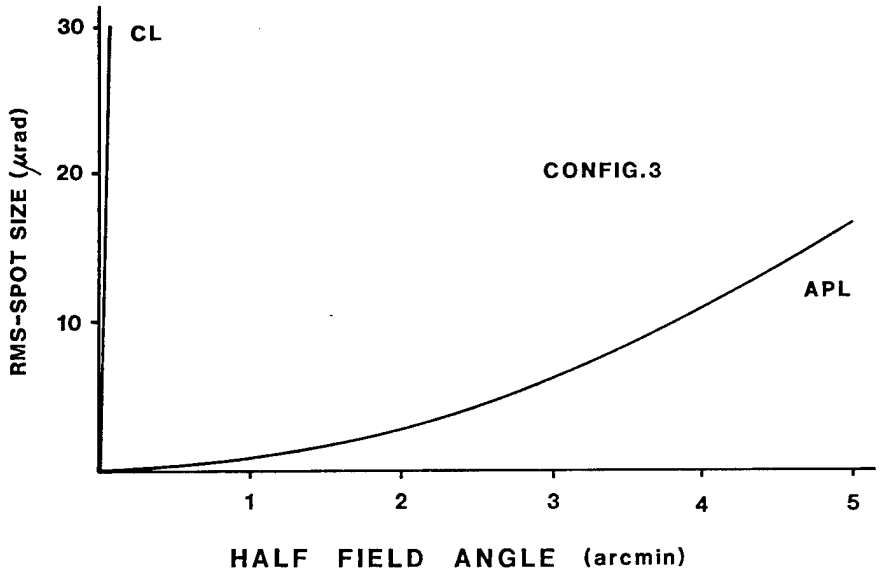


Figure 8. Performance comparison of classical (CL) and aplanatic versions (APL) of configuration 3.

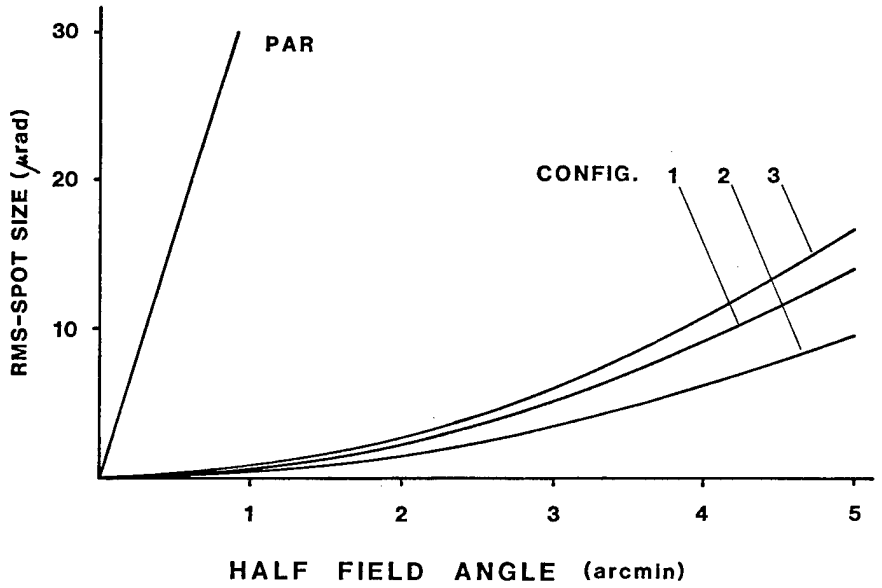


Figure 9. Performance comparisons of aplanatic two-mirror systems with an F/1 paraboloid.

References

1. NASA Advisory Council, Woods Hole Symposium, June 1979

2. NASA Astrophysics Projects Concept Summary, "100 Meter Thinned Aperture Telescope (TAI)", (in prep. at NSFC, 1980)

3. Korsch, D., Appl. Opt., 19, 499 (1980)

Design aspects of a laser gravitational wave detector in space

Rudolf Decher, Joseph L. Randall
NASA Marshall Space Flight Center, Marshall Space Flight Center, Alabama 35812
Peter L. Bender, James E. Faller
Joint Institute for Laboratory Astrophysics, Boulder, Colorado 80302

Abstract

Some of the optical and mechanical aspects of a proposed laser gravitational wave antenna in space are discussed briefly. The proposed concept consists of a free-mass antenna with the test masses separated by 10^6 km. A laser heterodyne technique is used to measure the distance change between test masses resulting from gravitational wave interaction. The proposed scheme appears to offer the necessary sensitivity to detect gravitational radiation from binary stars predicted by General Relativity Theory.

Detection of gravitational radiation is an important research goal in physics and astrophysics. Though Einstein's General Relativity Theory (1916) predicts the existence of gravitational waves, gravitational radiation has not yet been detected. Based on the successes of this theory so far, it is generally assumed that gravitational radiation does exist, and, in fact, several research groups are actively involved in the design of gravitational wave antennas. This paper outlines the concept of a space-borne gravitational wave antenna employing a laser heterodyne technique which, by exploiting the long base lines obtainable in space, appears to offer the necessary sensitivity to detect gravitational radiation from known sources.

Gravitation is the weakest interaction in Nature. The power radiated in gravitational waves as well as their effects on matter are extremely small and very difficult to detect. The design of any type of gravitational wave antenna represents a considerable challenge to science and technology.

General Relativity Theory predicts that gravitational radiation propagates with the speed of light and causes tidal forces which produce strain in matter transverse to the direction of propagation. The quadrupole nature of gravitational waves produces opposite strain in two mutually orthogonal directions, i.e., contraction and expansion. Gravitational waves are characterized by their frequency, dimensionless strain amplitude $h = \Delta L/L$, and polarization.

There are two classes of gravitational wave antennas: bar antennas and free-mass antennas. Gravitational waves interacting with a resonant bar antenna will excite extremely small free-mode vibrations which are detected by a sensitive transducer. Bar antennas are narrow-band detectors limited to the higher frequency range of gravitational radiation. Most of the ongoing development of gravitational wave antennas is in the area of cryogenic bar antennas.

The second class of antennas consists of free masses separated by a distance L . A gravitational wave causes a small change ΔL in the separation of the masses. The corresponding displacement is detected by electromagnetic signals transmitted between the masses. Free-mass antennas are broad-band detectors which take advantage of the fact that as one increases the separation L of the masses, the magnitude of the resulting displacement ΔL increases, therefore making it easier to detect. One approach that has been suggested involves precision microwave Doppler tracking of a spacecraft at interplanetary distance, with the earth and the spacecraft acting as free test masses.[1]

A much more sensitive space experiment to detect gravitational waves could use lasers to measure the difference in length of two arms of a large antenna. By keeping the arms very nearly equal in length, the sensitivity of the results to fluctuations in the laser frequency can be made small. The entire antenna would be in space and well away from the earth, so that perturbations due to disturbances on the earth's surface or propagation through the atmosphere are avoided.

The type of gravitational wave detector we will discuss uses laser heterodyne measurements to detect the difference in length of the two arms of the antenna. This approach has been proposed elsewhere and will be described briefly here.[2] The specific example considered involves arm lengths of 10^6 km. We will concentrate here on the engineering challenges presented by such large gravitational wave antennas in space.

An alternate approach using an equal-arm Michelson interferometer with arm lengths of 1 to 10 km has been discussed by Weiss and also appears feasible.[3] However, the heterodyne scheme permits longer antenna base lines to take advantage of the fact that the gravitational wave signal increases linearly with the antenna base line. Since major noise sources increase more slowly with the antenna length, large separation of the test masses should yield increased sensitivity.

Free-mass antennas in space permit the long base lines that are needed to detect very low-frequency gravitational radiation for which known sources exist. Presently the very low-frequency region appears to offer the best promise for detection of gravitational waves. Theoretical studies have been performed to identify possible astrophysical sources of gravitational radiation and to estimate intensity and frequency of gravitational waves arriving at earth.[4] Table I lists some potential sources of gravitational waves and the corresponding frequencies and strain amplitudes. While these estimates contain large uncertainties, they provide both direction and encouragement in the search for gravitational radiation.

The proposed laser heterodyne antenna aims at the detection of gravitational waves from rotating binary stars with known frequencies and locations. The most promising sources have been discussed by other authors.[5-8] The expected strain amplitudes $\Delta L/L$ range from approximately 0.4×10^{-21} to 6×10^{-21} and the gravitational wave periods from approximately 8 minutes to 193 minutes. In addition, the proposed antenna would be able to detect longer period pulses of gravitational radiation with strain amplitudes greater than 10^{-21}.

The geometry of the long base line laser heterodyne antenna is shown in Figure 1. The antenna consists of three satellites: the main satellite (S1) and the two subsatellites (S2 and S2). All satellites are in circular orbits around the sun, and their orbital periods are equal to one earth year. This will keep the antenna in the vicinity of the earth to provide easy communication and data transmission. The two subsatellites are in the same orbit, which is inclined by 3 degrees relative to the earth orbit plane. The main satellite is located above the subsatellite orbit to obtain a right-angle, equal-arm configuration. The base line between the main satellite and each subsatellite is 10^6 km.

Two-way laser signals are transmitted between the main satellite and each subsatellite. The antenna arms must be kept nearly equal in length to reduce the effect of laser frequency instability. With the assumed configuration the main satellite is approximately 700,000 km above the subsatellite orbit; i.e., its orbit is inclined by a very small amount relative to the subsatellite orbit. Consequently, the main satellite will move slowly toward the subsatellites. This downward motion can be compensated by occasional thrust maneuvers (once a day or every few days). The resulting motion must be at right angles to the line connecting the two subsatellites, in which case the Doppler effect is equal for both arms and will cancel out in the proposed heterodyne scheme. The trajectory of the main satellite consists of parabolic segments. During a day the velocity changes from 1.2 m/s up to 1.2 m/s down. Over an 8-hour period the motion would be only 3 km.

Each satellite will be equipped with a propulsion system for initial setup and occasional corrections to maintain the proper geometry. Standard microwave tracking can be used to accomplish position and velocity adjustments of the satellites.

The approach we would like to suggest to measure the separation is a two-way laser heterodyne distance measurement between each of the two spacecraft links, as shown in Figure 2. Each satellite is equipped with a local oscillator. A laser signal with phase $\omega_1 t + \phi_1(t)$ is sent from S1 to S2 and S3, where it is beat against local oscillators with the phase $\omega_2 t + \phi_2(t)$ and $\omega_3 t + \phi_3(t)$. The beat signals B_{21} and B_{31} are telemetered back to S1. Also, the local oscillator signals from S2 and S3 are transmitted back to S1, where beat signals B_{12} and B_{13} with respect to the laser oscillator at S1 are generated. Allowing for the times at which the signals are transmitted and received, we obtain:

$$B_{21} - B_{12} = \omega_1 \tau_{121} + \phi_1 (t) - \phi_1 (t - \tau_{121}).$$

A similar expression is found for $B_{31} - B_{13}$, with $\tau_{121}(t)$ and $\tau_{131}(t)$ the round-trip travel times for signals received back at S1 at time t. Finally forming $D=(B_{21}-B_{12}) - (B_{31}-B_{13})$, we obtain:

$$D = \omega_1 \{\tau_{121}(t) - \tau_{131}(t)\} + \phi_1 \{t - \tau_{131}(t)\} - \phi_1 \{t - \tau_{121}(t)\}$$

If the phase jitter in the laser over the difference in travel times for the two arms is sufficiently small, the final phase difference D is just 2π times the difference in length for the two arms, measured in terms of the wavelength corresponding to ω_1.

The preceding procedure has been used with microwaves for terrestrial distance measurements over single paths and for synchronization of remote oscillators.[9] To reduce the effect of laser frequency instability the two arms must be nearly equal in length. Laser frequency stabilities achieved to date over the rotating binary periods of minutes to hours are a few parts in 10^{13}. For the detection of gravitational waves from rotating binaries with strain amplitudes of the order of 10^{-21}, the arm lengths need to be equal to 1 m to avoid phase noise resulting from laser frequency instabilities.

The laser heterodyne communication system which we assume for signal-to-noise calculations consists of 10 mW beams at 0.633μm from cavity stabilized helium-neon lasers transmitted and received by 0.5 m diameter telescopes. The overall signal count rate would be approximately 10^7 photoelectrons per second, and roughly 10^4 seconds averaging time would be needed to detect a gravitational wave of 10^{-10} cm amplitude if no other noise sources were present. The necessary short-term phase stability in 0.633μm helium-neon lasers has been demonstrated in an experiment in which two lasers were mounted in the same frame in order to reduce acoustical noise in the beat frequency.[10]

The laser heterodyne communication system to achieve the precision pointing and tracking could be of the basic design developed during the 1960's for space laser communication.[11] Each terminal of the link would consist of an optical transceiver that has a laser transmitter and tracking system. Each transceiver would track the laser beacon of the other. To compensate for spacecraft motion, an optical image motion compensation technique could be employed. An alternate approach would be spacecraft body pointing similar to that used for the Space Telescope.

Many of the possible noise sources and perturbing forces have been discussed previously.[3] Perturbing forces on the test masses which have to be considered include solar radiation pressure, solar wind, gravity gradients, interplanetary dust, and cosmic rays. Shielding of the satellite is needed to reduce these perturbations below roughly 10^{-19} g, which is the equivalent level of accelerations expected from fast rotating binaries for a base line of 10^6 km.

Forces resulting from solar radiation pressure are considerably stronger than those caused by solar wind. But solar radiation pressure is quite constant over gravitational wave periods and can be eliminated by shielding. The most serious problem seems to be perturbations caused by solar wind. The acceleration resulting from solar wind pressure is quite variable over periods of minutes to hours.[12-13] To reduce the solar wind perturbations, a carefully designed shielding system employing drag-free system techniques is required.[14-16] The approximately 1.4×10^{-12} g acceleration expected for a 100 kg shield with 1 m^2 area can be estimated to have fluctuations of roughly 1×10^{-13} g in a cycle per day bandwidth at the frequencies of interest.

Several factors have to be considered in designing the shield for the satellites. To reduce disturbances introduced by the shield, the distance between the shield and the enclosed test mass should be made fairly large (e.g., to reduce gravity gradient and thermal radiation effects). How well the position of the shield with respect to the inner mass has to be maintained and how many additional factors would have to be controlled require further investigation.

Error sources in the laser/optical system include laser frequency instability (mentioned previously), Doppler shifts, amplitude noise in laser output power, and tracking and pointing errors which will cause amplitude modulation. Angular rotation of the telescopes will cause a change in the optical path length. Thermal stability is very critical to maintain the optical path length within the required limit. The two optical path lengths should be stable to 10^{-10} cm or better, which places stringent requirements on tolerance and alignment of the optical system. The proposed concept of a gravitational wave antenna in space requires much more analysis to evaluate error sources and to define details of the system. This is a challenging research area. Nevertheless, it appears that strain amplitudes from known sources could be detected within the constraints of available measurement technology.

The search for gravitational waves is a very important scientific endeavor. Failure to detect gravitational radiation would strongly affect our view of physics. Detection of gravitational radiation will open a new window to the Universe, initiating gravitational wave astronomy.

Table I. Some Possible Sources of Gravitational Radiation

Sources	Frequency or Period	Strain Amplitude $\Delta L/L$	Useful Antennas
Explosions in Quasars and Galatic Nuclei	10^5 to 10 s Burst	10^{-15} to 10^{-18}	Free-Mass Antennas in Space Spacecraft Doppler Tracking
Fast Binary Stars	10^4 to 10 s CW	$\sim 10^{-21}$	Free-Mass Antennas in Space Laser Heterodyne and Interferometer
Pulsars	1 - 100 Hz CW	10^{-23} - 10^{-27}	Free-Mass Antennas in Space
Black Holes Collapse of Stars Supernovae	100 Hz to kHz Burst	10^{-18} to 10^{-22}	Bar Antennas Ground-Based Laser Interferometer

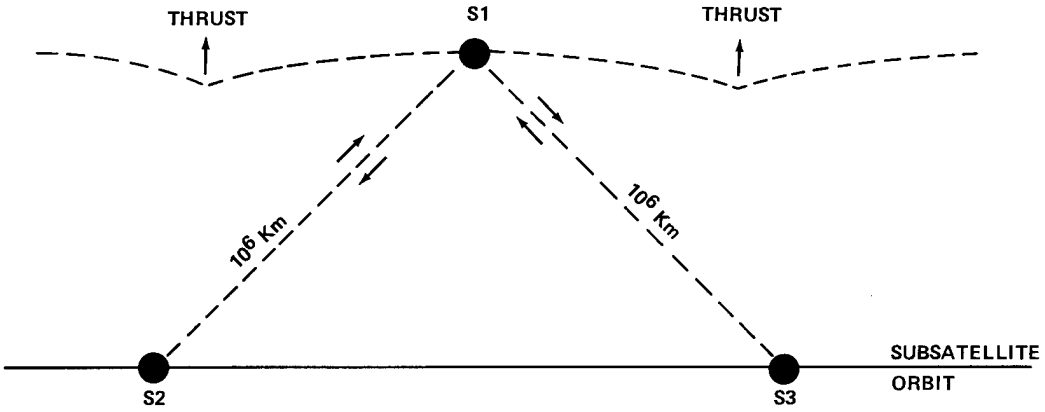


Figure 1. Gravitational wave antenna configuration

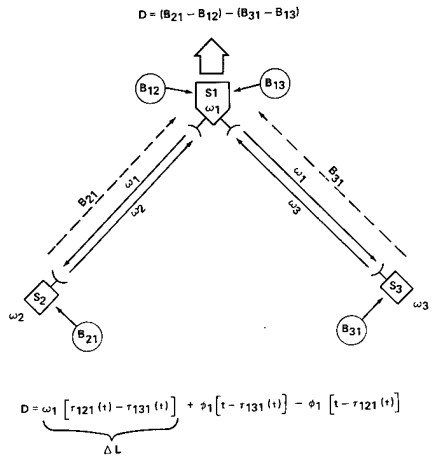


Figure 2. Signal heterodyne scheme

References

1. Estabrook, F. B.; Hellings, R. W.; Wahlquist, H. D.; and Wolff, R. S., in Sources of Gravitational Radiation, L. L. Smarr, ed. (Cambridge University Press, 1979).

2. Bender, P. L.; Faller, J. E.; Drever, R. W.; and Weiss, R., in preparation.

3. Weiss, R., in Sources of Gravitational Radiation, L. L. Smarr, ed. (Cambridge University Press, 1979).

4. For summary and references, see Sources of Gravitational Radiation (Discussion Session II), L. L. Smarr, ed. (Cambridge University Press, 1979).

5. Braginsky, V. B., Usp Fiz Nauk 86, 433 (1965) [Sov. Phys.--Usp 8, 513 (1966)].

6. Misner, C. .W.; Thorne, K. S.; and Wheeler, J. A., Gravitation (W. H. Freeman, San Francisco, 1973), p. 990.

7. Faulkner, J.; Flannery, B. P.; and Warner, Brian, Astrophys. J. 175, L79 (1972).

8. Warner, B., and Robinson, E. L., M.N.R.A.S. 159, 101 (1972).

9. Wood, L. E., and Thompson, M. C., Jr., Appl. Opt. 7, 1955 (1968).

10. Boersch, H.; Herziger, G.; and Lindner, H., Zeit. f. Physik 199, 349 (1967).

11. Precision Pointing and Tracking Techniques Development, Perkin-Elmer Corp., NASA Contract NAS8-20115 (1968).

12. Hundhausen, A. J., Solar Wind and Coronal Expansion (Springer-Verlag, Berlin-Heidelberg-New York, 1972).

13. Feldman, W. C.; Ashbridge, J. R.; Bame, S. J.; and Gosling, J. T., in The Solar Output and Its Variation (Colorado Assoc. Univ. Press, Boulder, in press).

14. Lange, B., AIAA J. 2, 1590 (1964).

15. Juillerat, R., in Experimental Gravitation, B. Bertotti, ed. (Academic Press, New York, 1974).

16. A. Satellite Freed of All But Gravitational Forces: TRIAD 1, AIAA J. Spacecraft and Rockets 11, 637 (1974).

AUTHOR INDEX

Preceding Page Blank

SUBJECT INDEX